# Attentional Processes in Implicit Sequence Learning

**Lee Rowland**

**University College London**

**Submitted to the University of London for the Degree of**

**Doctor of Philosophy in Psychology**

**May, 2006**





ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Recent conceptualisations of human learning and memory have drawn a distinction between conscious (explicit) and unconscious (implicit) processing modes (e.g., Clark & Squire, 1998). In line with this dichotomy, researchers have suggested that implicit learning is accomplished by automatic mechanisms that acquire information incidentally (Jiménez, 2003). Concordant with classical definitions of automaticity (Schneider & Shiffrin, 1977), the present thesis investigates whether implicit learning can be distinguished by its propensity to operate without placing demands on attentional resources and by its independence from selectional control. In contrast to previous studies, it was found that learning in the probabilistic serial reaction time (SRT) task is impaired by the presence of a secondary task designed to consume attentional resources (Experiment 1; cf. Jiménez & Méndez, 1999), and that selective attention during encoding is necessary for learning about an incidental to-be-ignored sequence (Experiments 6, 7 & 9; cf. Cock, Berry, & Buchner, 2002). Thus, these results do not support the existence of an automatic implicit learning system. Additionally, Experiment 1 presents evidence that SRT learning is conscious.

However, experimental procedures that interfered with input stages of SRT learning – by introducing irrelevant distractors into the display – revealed that implicit learning is highly resistant to disruption of the selection process (Experiments 2-5 & 9). Moreover, other experiments (Experiments 8 & 9) show that two complex probabilistic sequences can be learned simultaneously, which further indicates that such learning is robust in the presence of noisy input. Collectively, these findings are consistent with the view that implicit learning is subserved by a powerful incidental learning mechanism, yet, like explicit processes, requires attention and awareness to function optimally (St. John & Shanks, 1997).

For my Grandmother, who gives so much, and never expects anything in return.

Acknowledgements

My supervisor at University College London, Professor David Shanks, has been – and no doubt still is – a paragon of professionalism and integrity. As a scientist and a teacher he has been exactly what I have needed, and not a bit left over. Some psychologists have regarded David's stance on the issues surrounding learning and memory as unduly sceptical, but I have found him to show rigour in the place of zeal, and to exhibit fortitude in the face of controversy. The notion of unconscious mental processes remains one of the most seductive and cherished in modern theories of how the mind works, but it is vital for progress that our intellectual enquiries always appeal to the facts. I must give a nod to Professor Henry Plotkin who advised me during the twilight of his own career to track down Shanks if veracity and excellence was what I sought. Shortly thereafter, I asked David if he considers himself a behaviourist: "No," he said with a smile, "I am an empiricist." This is how I think of him, and it is his adherence to objectivity that I have striven to emulate.

Professor Nilli Lavie's work inspired my initial researches into attention, and the convenience of being able to run upstairs at short notice to discuss ideas with her – coupled with her enthusiasm – made for a very enjoyable working environment. Professor Jon Driver had a huge influence during my Masters at UCL, both on how I thought about the research process and in my understanding of attention. His intellectual fervour was infectious.

This research was supported by the United Kingdom Economic and Social Research Council, who awarded me a studentship.

# Contents

# Figures

# Tables

# Chapter 1

## Attention and Implicit Sequence Learning

Since Edward Thorndike's pioneering investigations into reinforcement awareness and the law of effect (e.g., Thorndike, *Human Learning*, 1931), psychologists have been intrigued by the notion of implicit learning, primarily because it may allow new insights into the functional roles of - and relationship between - cognition and consciousness. Common characterisations of implicit learning suggest that it differs from its explicit counterpart in at least three broad ways: first, it may take place incidentally, i.e. unintentionally, in the absence of hypothesis-testing strategies; second, it appears to occur automatically as a side-effect of task processing, and thus places minimal demands on attentional resources; and, third, its knowledge base is believed to be acquired independently of, and is subsequently inaccessible to, conscious awareness (for comprehensive reviews see Frensch & Rünger, 2003; Jiménez, 2003; Shanks, 2005). These claims have been elaborated and explored by a number of researchers offering a variety of experimental evidence as support: in addition to artificial grammar studies (e.g., Reber, 1967), learning tasks such as classical conditioning (e.g., Clark, Manns, & Squire, 2001), system-control tasks (Hayes & Broadbent, 1988) and serial reaction time tasks (e.g., Destrebecqz & Cleeremans, 2001) appear to validate the above conception of implicit learning.

Yet to some psychologists the notion of a separate learning mechanism governed by exclusive principles is implausible. In the case of awareness, it has proved particularly difficult to unequivocally support the thesis of 'unconscious' learning (e.g., Lovibond & Shanks, 2002; Shanks & St. John, 1994), in part due to the problem of defining awareness but also because of the elusiveness of finding compelling demonstrations of learning in conditions of null awareness (Merikle, 1994; Reingold & Merikle, 1988). In response to these difficulties surrounding the concept of awareness, some researchers have focused on the second characteristic described above and have placed emphasis on the idea that the implicit-explicit distinction should be recast on the grounds of attention. In this view, it is conjectured that implicit learning can occur with minimal, or zero, demands on general attentional resources and to that extent is automatic (Cleeremans, 1997; Curran & Keele, 1993; Frensch, 1998; Jiménez & Méndez, 1999; Jiménez, 2003). In contrast, on the basis of experimental evidence, other researchers have repudiated the attentional independence hypothesis of implicit learning

(e.g., Hsiao & Reber, 1998; Perruchet & Vinter, 2002; Shanks, 2003; Shanks & Channon, 2002; Shanks, Rowland, & Ranger, 2005), making this a currently contentious debate in the arena of cognitive psychology. The uncertainty about the role of attention in implicit learning is the focus of the present thesis.

## 1.1 Learning and Automaticity

What is the basis for the notion of a species of learning that operates differently from explicit learning with respect to attentional demands? That conscious attention plays a major role in explicit learning and memory is a robust and well-established empirical finding (e.g., Baddeley, Lewis, Eldridge, & Thompson, 1984; Broadbent, 1958; Cowan, 1995; Crabb & Dark, 1999; Dawson, Schell, Beers, & Kelly, 1982; Glucksberg & Cowan, 1970; Norman, 1969). Thus reasonably, a unique form of learning might depend on the existence of functionally or anatomically separate learning and memory systems, and the operation of cognitive processes that are liberated from conscious control. Indeed, over the last few decades the attempt to discover and characterise distinct memory systems has certainly been popular (e.g., Clark & Squire, 1998; Rolls, 2000; Shallice, 1988; Sherry & Schacter, 1987; Tulving, 1995; Tulving & Schacter, 1990), as too has the notion of the cognitive unconscious (Fodor, 1983; Kihlstrom, 1987; Reber, 1993). However, despite these approaches becoming increasingly in vogue, neither of them has stood up to scrutiny. The data upon which multiple-systems views have been based can often be explained more parsimoniously within the confines of a single-process model (e.g., Kinder & Shanks, 2001, 2003; Nosofsky & Zaki, 1998; Shanks & Perruchet, 2002). Similarly, the postulation of unconscious mental processes is often unnecessary to account for the facts about cognition (Perruchet and Vinter, 2002; Searle, 1992; Shanks & St. John, 1994).

Although the multiple systems view and the cognitive unconscious are doubtful, a genuine automatic process must, in adherence to classical definitions of automaticity, be distinguished by its independence of either selectional control or demands on central resources (Schneider & Shiffrin, 1977). This distinction between selection (i.e., selective attention) and resources (i.e., mental effort) is one traditionally made in the literature on attention (Johnston & Dark, 1986; Kahneman, 1973). Whilst it is unanimously accepted that attention functions as a selective mechanism aimed at focussing processing resources on some events at the expense of others (Driver, 2001), the concept of a finite pool of resources has been more difficult to pin down. For

instance, Neisser (1976) eschewed all notions of capacity limitations on human information processing; and Navon (1984) argued that it might be futile to demonstrate experimentally the existence of such resources. It has even been suggested that there might exist multiple independent pools of resources reserved for distinct processes (Navon & Gopher, 1979; Wickens, 1984) or modalities (Duncan, Martens, & Ward, 1997). Thus, attempting to demonstrate that some processes can proceed whilst making minimal, or zero, demands on these resources is likely to be problematic.

But perhaps more troubling is that convincing experimental verification of automatic processes has not been forthcoming. Stroop interference (Stroop, 1935) has been shown to be diluted by either the presence of distractors or by manipulating the focus of selective attention (Kahneman & Chajczyk, 1983; McCann, Remington, & Van-Selst, 2000), and thus even the quintessentially effortless process of word reading is not as automatic as previously thought. Moreover, recent accounts of emotion have demonstrated that the processing of angry faces, strongly believed to be automatic (Öhman, 2002), is subject to attentional constraints (Pessoa, 2005). Furthermore, despite appeal to the idea of automaticity, a recent survey of the literature has argued that it is not standing up as the robust phenomenon it was thought to be (Pashler, Johnston, & Ruthruff, 2001). Others too have challenged accounts of automatic processing (Styles, 1997) and learning (Cohen, Dunbar, & McClelland, 1990).

Curiously, most of the formal accounts of automaticity refer to the *development* of automatic processing, or skill learning (Logan, Taylor, & Etherton, 1999). It is not therefore clear why learning or memory could be expected to involve initial automatic encoding processes given that automatic processes usually develop after much effortful and time-consuming practice (Logan, 1992). Indeed, several existing studies demonstrate the critical role of processing resources in animal (Mackintosh, 1975), evaluative (Field & Moore, 2005) and classical (Dawson et al., 1982) conditioning. Thus, if the atomic products of learning require attention during encoding, and severe doubts remain over the notion of automatic processes in the absence of effortful practice, then the claim that implicit learning makes no capacity demands is to be seriously questioned.

Presumably, implicit learning has been singled out as a candidate automatic process because it appears to occur incidentally as a side-effect of goal-relevant processing (Cleeremans, 1993, 1997; Cleeremans & Jiménez, 1998; Jiménez & Méndez, 1999, 2001; Jiménez, Méndez, & Cleeremans, 1996; Reber, 1967, 1993). From this

perspective, researchers have speculated that attentional involvement may be minimal in a form of learning that does not seem to require the usual level of engagement common to intentional explicit learning. This idea is attractive because such a system would allow for the effortless acquisition of powerful cognitive abilities, such as language or complex skills, driven predominantly by stimulus inputs.

## 1.2 The Paradigm of Sequence Learning

Originally devised by Nissen and Bullemer (1987), the serial reaction time task (henceforth, SRT) has proved to be a flexible methodology for investigating implicit learning, and consequently it has been employed, often with modifications, by many researchers in the field of implicit learning (Cleeremans & McClelland, 1991; Cohen, Ivry, & Keele, 1990; Destrebecqz & Cleeremans, 2001; Frensch, Buchner, & Lin, 1994; Frensch, Lin, & Buchner, 1998; Jiménez & Méndez, 1999; Jiménez, Méndez, & Cleeremans, 1996; Perruchet & Amorim, 1992; Reber & Squire, 1998; Reed & Johnson, 1994; Schvaneveldt & Gomez, 1996; Shanks & Channon, 2002; Shanks & Johnstone, 1999; Shanks, Wilkinson, & Channon, 2003; Stadler, 1995; Willingham, Greenberg & Thomas, 1997). In this task the display typically comprises four or more horizontally aligned screen locations, and on each trial participants are required to respond to the location of a target stimulus by pressing a key as quickly as possible. Crucially, the order of target locations follows a repeating sequence. As participants become sensitive (perhaps unconsciously) to the underlying structure of the sequence, successive locations become primed resulting in speeded responses to structured versus random trials. Because this sequence knowledge has sometimes been difficult for participants to express verbally, the type of learning in (at least some of) these studies has been classified as "implicit".

The SRT task is appealing because it disguises the fact that participants are in a learning environment. The sequential structure is incidental to the target response task, which may reduce the possibility of participants attempting to consciously search for reasons underlying their improved performance (but see Frensch, Haider, Rünger, Neugebauer, Voigt, & Werg, 2003, for an alternative view). A further desirable feature of the SRT task is that people learn a low-level perceptual-motor skill. Often the rules required to distinguish between grammatical and ungrammatical items in artificial grammar learning may be quite complex, and moreover, it is unclear whether performance draws on rule knowledge (Perruchet & Vinter, 2002; Shanks, Johnstone, &

Kinder, 2002), thus making it difficult to assess what contributes to learning. Sequence learning does not suffer from these problems because only sequence knowledge can facilitate responses to structured trials relative to random trials.

### 1.2.1 Dual-Task Studies of Sequence Learning

To investigate the role of attention in implicit sequence learning, researchers developed dual-task versions of the SRT task with the aim of dividing attentional resources between the response task and a concurrent attention-demanding task. Undoubtedly, the underlying assumption about dual-tasking is that concurrent effortful tasks compete for a limited reservoir of resources and that sharing this capacity would in some way impair processing of the primary task (Navon, 1984; Pashler, 2000). Some studies report that learning under such dual-task conditions is equivalent to the degree of learning obtained on the SRT task alone (Cleeremans & Jiménez, 1998; Curran & Keele, 1993; Frensch, et al., 1998; Frensch, Wenke, & Rünger, 1999; Heuer & Schmidtke, 1996; Jiménez & Méndez , 1999, 2001; McDowall, Lustig, & Parkin, 1995; Schvaneveldt & Gomez, 1998; Seger, 1997; Stadler, 1995). These results have been taken to support the proposal that implicit learning proceeds automatically.

The study by Frensch et al. (1998) provides a good example. In their experiments, participants were trained for several hundred trials on 9-location sequences. The major independent variable was the inclusion of a tone-counting secondary task during the training stages. In this task, participants were required to keep a running count of the high- but not the low-pitched tones that were presented during the inter-target intervals of the SRT task. All groups performed the tone-counting task, but crucially, the most extreme 'dual-task' group trained predominantly under dual-task conditions, whereas other groups had considerably less exposure to the tone-counting task. For all groups, the tones were removed for the testing stage, and the training sequence was replaced by a quasi-random sequence for two blocks, before reverting to the structured sequence. The response time difference between structured and random sequences provided a measure of learning. The results were clear: transfer scores were almost identical across groups, suggesting that the demanding tone-counting task does not interfere with sequence learning.

However, Shanks and Channon (2002) identified inherent weaknesses in the Frensch et al. methodology. For example, Shanks and Channon argued that by intermixing dual- and single-task training conditions for both groups, the study

(Experiments 1a and 1b) did not provide a powerful test of the hypothesis that sequence learning may proceed independently of attentional demands, relative to a comparison between groups that only received single- or dual-task training.

Another problem was that Frensch et al. included in their analysis participants who had performed rather poorly at tone counting. These participants might have been allocating more attention to the SRT task than was ideal and thus learning under dual-task conditions cannot be said to have occurred exclusively in the absence of attention.

Finally, the sequences used by French et al. possessed some undesirable properties, such as the presence (in the training sequences) and the absence (in the test sequences) of reversals. Instead of acquiring sequence knowledge, participants may have learned an abstract feature of the training sequence that led to inflated transfer scores, thus clouding the issue regarding the effect of the secondary task.

To counter these limitations, Shanks and Channon (2002) modified the Frensch et al. procedure and attempted to replicate their findings. Firstly, participants were trained only under dual- or single-task conditions. They were subsequently tested under either dual- or single-task conditions, but the between-participants design ensured that both training groups were tested under both test conditions. The dual-task training group performed a tone-counting secondary task during the SRT task. Participants reported their count at the end of the block of trials and those with accuracy less than 90% were excluded from the analysis. In contrast to the Frensch et al. (1998) experiment, Shanks and Channon ensured that the training and the quasi-random test sequence were structurally identical, and they improved the quality of their sequences by adopting the Reed and Johnson (1994) structures (see *Method* section of Experiment 1 herein for details). Contrary to Frensch et al.'s observations, Shanks and Channon obtained significantly greater transfer scores for the single-task group than the dual-task group. This demonstrates that sequence learning can be impaired by a secondary task. Furthermore, the same pattern of results was observed under conditions of dual-task testing (although overall, transfer scores were lower), which suggests that expression of sequence knowledge is not affected by the test conditions (cf. the "suppression hypothesis" of Frensch, 1998).

## 1.3 Selective Attention and Implicit Learning

The preceding section cast doubts over the attentional independence claim of implicit learning; but what about the other kind of automaticity, that of selectional

control? There have been far fewer investigations into the role of selective attention in implicit learning. In the case of a standard SRT task, selective attention seems inevitable: correct responses are based on directing attention to the spatial locations of targets. The question remains, however, whether manipulations of selective attention would affect sequence learning?

Jiménez, Méndez and Lorda (1993) developed a simple test to try to establish whether selective attention is necessary for implicit sequence learning. Their SRT task comprised a structured sequence of locations coupled with a predictive relationship between stimulus colour and location. Participants merely had to respond to the location of the stimulus regardless of its colour. Despite the locations following a complex sequence of ten locations, and the successive location being predicted by the colour of the square on the current trial, participants learned about the complex sequence of locations yet very little about the much easier predictive relationship between colours and locations.

These results at first glance appear to suggest that lack of selective attention to a predictive dimension prevents learning about such contingencies. There are however alternative interpretations of these findings. Jiménez and Méndez (1999) argued that the associative process could have been prevented or hindered by the dimensional switch between the current colour predicting the successive location, which is potentially more disrupting and harder to learn than a continuous predictive relationship among locations. Furthermore, there was no comparison between a group trained on the above task and one that required participants to pay attention to the colour dimension. Such a design would separate out the effects of selective attention and dimensional switch. This is exactly what Jiménez and Méndez (1999) did in a later series of experiments (see below).

Using a different implicit learning task, Chun and Jiang (1998, 1999, 2003; Jiang & Chun, 2001; see Jiang & Chun, 2003, for a review of their work) have demonstrated that implicit learning and attention enjoy reciprocal influences. In their contextual cueing task, participants respond to the presence or absence of a target item located within a display of distractor items (visual search). Importantly, this differs from traditional visual search tasks because some distractor-target layouts repeat across blocks of trials. Jiang and Chun have discovered that participants learn to respond faster to target items that are predicted by the spatial layout of the display. Because learning in these studies appears to be implicit (as participants do not seem to be aware of the

repetition of the displays), it has been argued that implicit learning is involved in the allocation of spatial attention (see also Lambert, 2003). Moreover, spatial attention also appears to facilitate implicit learning. To establish this, Jiang and Chun (2001) manipulated where attention was focussed during the contextual cuing task. They presented displays simultaneously containing items in two sets of colours, and informed participants to attend to one colour only as this would assist them in the target identification task. A contextual cueing effect was only obtained when the attended colour was repeated across trials, and not when the ignored colour was repeated. Thus, even though the ignored set was predictive of target location, participants learned nothing about this dimension because they did not attend to it.

The evidence that implicit learning is dependent on selective attention is consistent with similar findings in the field of implicit memory (Mulligan & Brown, 2003), and it seems therefore that attention to items is a fundamental requirement for access to basic learning mechanisms. However, there is some evidence that militates against this view. Cock, Berry, and Buchner (2002) trained participants on a novel version of the SRT task whereby both a target sequence and an irrelevant sequence were presented simultaneously. Although participants were required to respond only to the target sequence, and to ignore the irrelevant sequence, significant negative priming effects were obtained across all three experiments, and these effects extended beyond mere individual events to include sequential features of the ignored stimuli. These results indicate that people can learn about a secondary sequence without the requirement to actively attend to it.

In essence, there seems to be a strong indication that selective attention is required for implicit learning. However, it is too soon to say for sure given that only a handful of studies exist, and that there is some evidence that ignored sequences can be learned. Certainly, more work needs to be done to confirm the precise role that selective attention plays in implicit learning.

## 1.4 Multiple Sequence Learning

The study by Cock et al. (2002) suggests that two sequences can be learned simultaneously. This proposal is consistent with the idea that implicit learning is automatic with respect to its independence from selective attention. But it also entices one to revisit the question about the role of mental effort in sequence learning. Indisputably, if people can learn two sequences concurrently, then this deserves

consideration as a possible demonstration of implicit learning proceeding independently of the demands on central resources.

This approach has not been exploited to anything like the degree of the dual-task procedure. In fact, very few studies address this issue directly. In one of these, Mayr (1996) designed an experiment to investigate whether an object sequence and a spatial sequence can be learned simultaneously. Four widely-separated screen locations were presented and a stimulus appeared at successive locations in a structured fashion. However, participants did not respond to the locations of targets, but on the basis of colour-shape conjunctions (e.g., a black square). This target identity comprised a second predictive dimension, such that the order of target-response mappings was structured (i.e., the colour-shape conjunctions followed a pattern). This object sequence was learned, but the key finding was that although the sequence of locations was independent of response class, there was evidence to suggest that spatial sequence learning had been accomplished even in the context of the collateral object discrimination task. To determine whether learning the secondary (location) sequence impacted on the primary (object) sequence, Mayr compared sequence learning under dual conditions (the two structured sequences presented simultaneously) with learning under single conditions (one structured sequence and one random sequence). The results showed that learning two sequences simultaneously did not compromise primary sequence learning, and led Mayr to conclude that people possess two independent learning systems, one for learning about objects and another for learning about spatial structures.

However, the problem with this conclusion is that it is not known whether people devoted resources to attempting to learn the random sequence. Although implicit learning mechanisms may detect a secondary (random) sequence in the absence of intentional efforts to learn, it does not follow that processing of this sequence occurs independently of attention. Ideally, Mayr should have included single sequence groups who responded to one sequence whilst the other dimension did not vary. Mayr's study is thus a weak test of the hypothesis that two sequences can be learned without cost relative to learning a single sequence alone, and consequently, the results should be viewed with caution.

## 1.5 Testing the Limits of Automatic Learning

Thus far, three approaches to investigating whether implicit sequence learning is an automatic process have been identified: dual-task studies; learning of ignored sequences; and multiple sequence learning. The contradictory evidence in each case has obviated the drawing of any firm conclusions.

However, a unique contribution to these questions has been made by integrating all three approaches in a single study. In a sophisticated and elegant design, Jiménez and Méndez (1999) manipulated mental effort and selective attention, and in so doing, were able to provide new evidence that two sequences can be learned simultaneously without cost to learning the primary sequence. In their study, Jiménez and Méndez trained participants on 31,000 trials of a sequence learning task, under either single- or dual-task conditions. They used a novel method for investigating the role of selective attention by arranging a predictive dimension between the shape of the targets upon which the secondary task was performed and the locations in the primary sequence. The shape-location sequence was simpler than the primary location sequence, and was constructed so that the effects on response times to the primary sequence could be analysed separately for the spatial dimension and the shape dimension.

In addition to the design, Jiménez and Méndez (1999) incorporated two methodological features that avoided the problems that befell previous work in this area. First, Jiménez and Méndez trained participants on probabilistic (noisy) sequences, which have been an important enhancement to SRT learning in recent years (Cleeremans & Jiménez, 1998; Cleeremans & McClelland, 1991; Schvaneveldt & Gomez, 1998). Under this preparation, the structure of the generating sequence is disrupted on a trail-by-trial basis by interspersing random elements with sequenced transitions. The difference in reaction time to structured versus noise trials allows for a continuous on-line measure of learning, and thus the reliance upon a distinct test phase can be dispensed with (a problem highlighted by Shanks & St. John, 1994). A second attraction of probabilistic sequences is that the random events potentially disrupt any attempt to corroborate a consciously registered relationship (Cleeremans & Jiménez, 1998), which makes them more likely to engage automatic learning mechanisms, should they exist (Jiménez, 2003). Additionally, noisy sequences provide a learning environment that exhibits similarities to the natural world, whereas it is rarely the case that learning is as structured and as uniform as learning a deterministic sequence.

Another feature introduced by Jiménez and Méndez (1999) was an alternative secondary task. Recent work on the tone-counting procedure has undermined its suitability as a concurrent attention-demanding task (Rah, Reber, &, Hsiao, 2000), showing that it is detrimental to implicit learning not because it depletes attentional resources, but because it introduces a set of co-occurrences that interfere with the structure of the SRT task. Rah et al. suggest that during sequence learning participants routinely scan for patterns of covariation, and that the introduction of a secondary task with no predictive value dramatically reduces the ability to detect such patterns. Thus, the concurrent task does not compete for resources as such, but disrupts the manner in which the sequence information is processed (for similar views see Heuer & Schmidtke, 1997; Jiménez & Vázquez, 2005; Stadler, 1995). Interestingly, if the tone task is structured such that it is more easily integrated with the target sequence, then the interference from this 'dual-task' is eliminated (Jiménez & Vazquez, 2005; Schmidtke & Heuer, 1997).

To overcome these problems, Jiménez and Méndez (1999) instead presented one of four symbols (e.g., 'x', '?', '!', '*') as the SRT target on each trial. During each dual-task block, participants were required to keep a running total of the number of times that two of the symbols had been presented. Jiménez and Méndez surmised that this secondary task is a purer manipulation of attention because it is performed on the stimuli that are the targets in the SRT task. Thus, whilst it might still involve a small degree of stimulus processing (in order to characterise a symbol as a count target), this should be only mildly disruptive in comparison with that of tone counting.

Jiménez and Méndez obtained three key results: first, learning the primary sequence was equal for the single and dual-task groups; second, in conjunction with learning the primary sequence, participants could also simultaneously learn about the secondary predictive dimension between shapes and locations; but, third, this only occurred in the dual-task group who were required to selectively attend to the shapes for counting. Collectively, these results suggest that sequence learning can proceed automatically, provided that participants pay selective attention to the predictive dimensions. In this sense, implicit learning of probabilistic sequences is automatic to the extent that it does not rely on additional attentional resources above those needed for processing the target stimuli (selective attention), which was one of the criteria for automaticity set out by Schneider and Shiffrin (1977).

Although the main findings are extremely convincing at first sight, there are several reasons to be cautious about generalising them. For instance, the extremely long training duration might have allowed participants to automatise the secondary task, and therefore, its demands on attention would not have been sustained throughout the experiment, leading to similar degrees of attention to the primary task for both the single- and dual-task groups. A second concern is that there were two concurrent predictors of target location, namely, locations and target shape on the previous trial. It is possible that learning in the dual-task condition was impaired, but that this was offset by the advantage of having target shape information that in some way augmented location sequence learning.

It is not therefore clear whether Jiménez and Méndez would have obtained the same results had they not had shape-location regularities, and had they not provided participants with the opportunity of automatising the secondary task.

## 1.6 Attention

The discussion thus far has treated attention as a simple concept with two components: a selective function, and a pool of processing resources. But attention is actually a very complex phenomenon that is only just beginning to be understood. In a comprehensive review, Driver (2001) noted that there are potentially many specific processes comprising "attention", and as the ensuing discussion will make clear, attentional processes cannot be easily defined by recourse to an elementary selection/resources dichotomy. A more sophisticated perspective on the nature of attention is likely to be highly informative when addressing questions about the role of attention in implicit learning, and encouragingly there has been a growing trend towards this in recent years (Jiang and Chun, 2003; see also other articles in the volume edited by Jiménez, 2003). The discussion will focus on visual attention because this is most relevant to sequence learning and is the most thoroughly investigated by attention researchers.

In its broadest sense, attention is the process by which the brain deals with the volume and complexity of sensorial input in a manner that serves to reduce the information flow upon which goal-directed behaviour operates. Restrictions on the amount of material that can be processed to a semantic or conscious level (because neural capacity is limited) impel the brain to select some objects for detailed representation at the expense of others. This has led to the proposal that competition for

neural resources is biased (Desimone & Duncan, 1996; Kastner & Ungerleider, 2000; Yantis, 2005). This competition is biased by both bottom-up stimulus-driven neural mechanisms that code stimulus features and by top-down control mechanisms that select objects that are relevant to goal-directed behaviour. In the absence of top-down influences, bottom-up biases favour objects that stand out in a scene.

But critical to the guidance of selective attention are the influences of top-down components (Kastner & Ungerleider, 2000; Maruff, Danckert, Camplin, & Currie, 1999; Remington & Folk, 2001). According to Desimone and Duncan (1996), working memory circuits in prefrontal cortex set a template for objects or locations to which to attend and these signals modulate activity in lower visual areas, thus biasing competition towards goal-relevant information. In fact, frontal functions such as working memory are involved in the control of many operations relevant to current behaviour, such as directing the locus of selective attention (de Fockert, Rees, Frith, & Lavie, 2001), the setting of task priorities (Baddeley, 1996), the selection and monitoring of responses (Humphreys & Samson, 2004), the inhibition of irrelevant information (Beer, Shimamura, & Knight, 2004), the maintenance of goal-states (Davidson, Jackson, & Kalin, 2000), and the resolution of conflict (Bush, Luu, & Posner, 2000). Moreover, evidence is mounting that top-down influences affect selective attention at multiple levels of stimulus processing (Remington & Folk, 2001), with results now indicating that performing demanding tasks can abolish neural activations to concurrently presented moving stimuli (Rees, Frith, & Lavie, 1997) and words (Rees, Russell, Frith, & Driver, 1999). Thus, the evidence suggests that perception is highly dependent on the availability of attention, with possibly only very rudimentary stages of perception free from its effects (Lachter, Forster, & Ruthruff, 2004; Treisman, Vieira, & Hayes, 1992). It is the dynamic interplay between control and perceptual mechanisms of attention that gives selected objects access to memory and motor systems for guiding behaviour. It follows that successful deployment of attention is indispensable for forming the contents of conscious visual experience (Baars, 2002).

The ideas described above offer a distinction that has been recognised in several models of attention, that between central (modality-independent) and peripheral (modality-specific) aspects of attention (e.g., Duncan, et al., 1997; Johnston, McCann, & Remington, 1995). There seems to be a convergence between researchers on at least one key idea: that selection mechanisms and control mechanisms are separate yet

interdependent functions of the human attentional system, each with their own capacity limits. In this sense, selective attention is affected by constraints at the levels of both perceptual representation and cognitive control, with each modulating processing of the other, as the next section shows.

### 1.6.1 Selective Attention Under Load

Maintaining attention on selected objects is often vital for successful performance in many tasks, yet frequently, other objects can distract us from what we are doing. It is therefore essential that irrelevant or interfering distractors be prevented from gaining control of behaviour, especially in situations where they are likely to be detrimental to performance. One enduring question is whether irrelevant objects are filtered during early stages of visual perception, or whether interference from irrelevant items is suppressed at the level of response output (e.g., Lachter et al., 2004).

There is strong evidence that the extent of irrelevant distractor processing is highly dependent on the type of load under which attention must operate (Lavie, Hirst, de Fockert, & Viding, 2004). Under conditions of low perceptual/sensory load (few relevant input items), residual processing capacity not consumed by goal-relevant tasks is available for processing irrelevant items in a scene. However, under conditions of high perceptual load (many relevant input items), there is very little spare processing capacity for irrelevant items because it is consumed by the difficulty of the central task. Lavie and colleagues' position regarding perceptual load can be generalised by saying that the more one must attend to successfully completing a task, the less likely it is that irrelevant information will be afforded processing resources.

There is however an important proviso to this latter assertion. The efficiency with which irrelevant distractors can be filtered depends on the availability of cognitive control functions to maintain the focus of attention (see also Lavie & de Fockert, 2005). Under conditions of high working memory load, for example, it becomes more difficult to maintain the locus of attention and more difficult to suppress the interfering effects of low-priority or irrelevant stimuli. Importantly, the degree to which working memory is involved in preventing distractor processing may depend critically on the degree of competition between target and distractors, only being necessary if distractors are salient (Lavie & de Fockert, 2005).

In support of Lavie's (e.g., 2005) theory there is a wealth of behavioural and imaging data. Some of the imaging effects are very compelling, showing reduced neural

responses in primary visual cortex to irrelevant distractors under conditions of high perceptual load, and increased neural responses in visual cortex to distractors under conditions of high working-memory load (de Fockert et al., 2001).

In many respects Lavie's work on attentional load appears to offer a resolution to the longstanding debate about early- versus late-selection models of attention: whether unattended items are filtered early or late in the perceptual stream depends critically on the degree and type of load under which people perform.

## 1.7 Interactions Between Attention and Implicit Sequence Learning

The foregoing discussion makes two key points. First, that attention is mediated by a number of independent mechanisms that have specific roles in processing information relevant to current behaviour. And second, that it might be inappropriate to think of selection and resources as independent aspects of the attention apparatus, when in fact they are inextricably linked. In the context of implicit learning, it could prove very informative to apply current knowledge in the psychology of attention to further investigate the complex interactions between attention and learning.

One widely-accepted view about implicit learning is that it depends on a non-selective processing system that routinely scans the environment for patterns of covariation (Cleeremans, 1993; Heuer & Schmidtke, 1996; Howard & Howard, 2002; Rah et al., 2000; Reber, 1993). If this is correct, then implicit learning should be highly reliant on control functions, such as working memory, in order to segregate and cohere associated input streams. Importantly, neural networks in prefrontal cortex are believed to subserve both working memory (Pessoa & Ungerleider, 2004) and sequence learning (Hazeltine & Ivry, 2003), so it seems reasonable to suspect that loading working memory will impair implicit sequence learning. Conversely, implicit learning mechanisms might be resistant to load on input (perceptual selection) mechanisms in order for many dimensions of the sensory input to be scanned for statistical regularity.

By considering attention and implicit learning in this manner, it will be possible to move beyond the basic questions about whether implicit learning is dependent on resources or selection, and begin to tease apart the intricate ways by which attention and learning systems interact.

# Chapter 2

## Attentional Load and Implicit Sequence Learning

The experiment reported in this chapter was designed to provide new evidence about whether implicit sequence learning operates independently of the demands on central attentional resources. The Jiménez and Méndez (1999) study referred to in Chapter 1 appears to strongly support this idea, but there are several features of their experiments that may complicate their results.

First, Jiménez and Méndez trained participants for extensively longer on the SRT task than has often been the case in previous work on SRT learning. Participants underwent 10 training sessions (carried out at the rate of 2 per day) each composed of 20 blocks of 155 trials. Therefore participants had a total of 31,000 learning trials compared with other sequence learning experiments that tend to use between 1,000-2,000 (e.g., Frensch et al., 1998; Shanks & Channon, 2002). Jiménez and Méndez analyzed their data in blocks of 3,100 trials – hence the very first data point on their learning curves is considerably longer than the entire training stage of many previous experiments. It is possible that with such substantial training participants simply became very efficient at task sharing; the secondary task may have become automatic therefore allowing attentional resources to be allocated to sequence learning. Indeed, whereas the overall RT difference between single- and dual-task groups in the Shanks et al. (2005) experiments was never less than 100 milliseconds, in Jiménez and Méndez' study it was only about 40 milliseconds by the third of their blocks of 3,100 trials and was eliminated entirely by the end of training. In one of their experiments (Experiment 3), Jiménez and Méndez did require participants to count a different pair of symbols in each block, but that still means that the same targets were counted across the whole of the first block of 3,100 trials. Moreover, changing the to-be-counted targets does not eliminate the possibility that target-nonspecific counting operations could have become automatized. It is also problematic that in their studies (similar to Experiments 1 and 2 in Shanks et al., 2005) the response-stimulus interval (RSI) was 240 milliseconds which provides a considerable time period in which the participant may update his or her symbol count before the appearance of the next target. Indeed Jiménez and Méndez themselves favoured the conclusion (p. 255) that "participants may have learned over training about how to best interleave the different requirements of their tasks to perform

them optimally." But of course if the two tasks were reconfigured in such a way as to allow efficient timesharing, it is no longer obvious that the results speak to the issue of the attentional demands of implicit learning. By interleaving the counting and target-response tasks, participants were ensuring that they did not compete for a common resource.

A second reason for caution in interpreting Jiménez and Méndez's results is that they arranged predictive relationships not only between locations, but also between the shape occurring on one trial and the location of the target on the next. They found (see also Jiménez & Méndez, 2001) that dual-task participants learned both of these sets of contingencies. However, this means that attending to the shapes provided participants with information (over and above their location) which was relevant to learning the sequence of locations. Consider a typical location sequence chunk such as 3-1-2-4-1 which might have occurred regularly for a given participant. If we denote the identity of the shapes as A-D, then this sequence might actually have been realised as 3D-1A-2B-4D-1C where the initial 3D, for instance, means shape D appearing at location 3 (note that the same shape precedes both trials on which the target appears at location 1). It is obvious that, from the point of view of elementary learning processes, this regularity of shapes and locations creates a sequence-learning problem that is informationally quite different from (indeed richer than) the sequence presented to a single-task participant who is completely ignoring the shapes and hence only apprehending a sequence of locations. Examples of augmentation or potentiation of the learning of one contingency by another simultaneous one (Batsell, 2000; Durlach & Rescorla, 1980) raise the possibility that dual-task participants were only able to match the learning of single-task ones because the negative costs (in terms of an attention decrement on sequence learning) of having to perform the secondary tasks were offset by the advantage of having target shape information which in some way 'supported' location sequence learning. Put differently, it is not at all obvious that Jiménez and Méndez would have obtained the same results if there had been no shape-location regularities and if the secondary task had provided no possible support for learning the primary sequence.

In response to the above concerns, Shanks, Rowland, and Ranger (2005) conducted a sequence learning study that, in concordance with Jiménez and Méndez (1999), incorporated probabilistic sequences and a symbol-counting secondary task, but in contrast to Jiménez and Méndez, used shorter training lengths and did not arrange a predictive relationship between the primary and secondary tasks. In Experiments 1 and

2 of the Shanks et al. study, participants were trained on probabilistic sequences for 9 (Experiment 1) or 14 (Experiment 2) blocks of 100 trials each under either single- or dual-task (symbol-counting) conditions, and then tested on a final block under single-task conditions (i.e., the secondary task was removed for the dual group). In both these experiments the single-task groups obtained learning scores that were considerably (and significantly) higher than those of the dual groups (nearly 30 milliseconds more in both cases). These results conflict with those of Jiménez and Méndez by demonstrating that sequence learning is impaired by the requirement to perform an attention-demanding secondary task.

## 2.1 Experiment 1

The present experiment (reported in Shanks et al., 2005, Experiment 3) is very similar to the studies described above but sought to avoid two possible criticisms of the previous two experiments. First, it might be argued that testing all participants under single-task conditions is unfair on the dual-task group in that for this group – but not for the single-task learning group – there is a change of conditions between the learning and test phases. Whereas the single-task group performs the same task in the test block as in the training blocks, the dual-task group has to adjust to a different set of conditions (i.e., the absence of the secondary task) and this may introduce a confounding factor. Perhaps learning is normal under dual-task conditions, but participants find difficulty in fully expressing their sequence knowledge under the transfer conditions. To counter this possibility, participants were tested under both single- and dual-task conditions in the present experiment.

Secondly, it might be objected that – while training in Jiménez and Méndez was too long - the length of training in the first two experiments in Shanks et al. (2005) was inadequate to allow the development of attention-independent implicit learning. Therefore the training stage is extended to 4,000 SRT trials.

On blocks 10, 20, 30, and 40 of the experiment participants in the single-task group were tested under dual-task conditions. Dual-task participants performed the secondary task on all blocks except block 40 where they were tested under single-task conditions. It was thus possible to compare sequence knowledge under dual-task testing conditions by analysing: (i) RTs on blocks 10, 20, 30, and (ii) RTs of the single-task group on block 40 with those of the dual-task group on block 39. In addition, sequence

knowledge can be compared under single-task testing conditions by analysing RTs of the single-task group on block 39 with those of the dual-task group on block 40.

Two other features of the present experiment also contributed to its usefulness in providing new information.

First, the response-stimulus interval was reduced to 0 milliseconds in this experiment. The reason for doing this is that the longer interval (250 msec) in the previous experiments may have provided an opportunity for participants to interleave the different requirements of the localisation and counting tasks so as to reduce competition between them for central resources. For instance, this interval provides an opportunity for participants to decide whether the previous symbol needs counting and to update their symbol count if necessary. Setting this interval at 0 milliseconds should eliminate this possibility. In other words the previous experiments, by allowing for such interleaving, may have underestimated the cost of the secondary task on sequence learning as they may have reduced the concurrent demands of the two tasks for a central attentional bottleneck.

Second, as has previously been used in other sequence learning studies (e.g., Destrebecqz & Cleeremans, 2001; Perruchet & Amorim, 1992; Shanks & Channon, 2002; Shanks & Johnstone, 1999; Shanks et al., 2005; Shanks, Wilkinson, & Channon, 2003) participants will perform a post-SRT test to assess their sequence awareness. Although this knowledge is not of primary interest to the present study, it may prove useful for two reasons. First, it would be illuminating to provide some data that speaks directly to the issue of whether probabilistic sequence knowledge is available to consciousness, and moreover, if it is, then does training under dual-task conditions reduce the level of explicit knowledge? And second, Shanks and Channon (2002) reported that in their experiments a participant's level of explicit knowledge about the sequence was not responsible for the reduced degree of learning observed for the group trained under dual-task conditions, whereas Frensch et al. (1998) claim that sequence learning can be independent of attentional demands only if participants lack explicit knowledge of the sequence, i.e., if learning is implicit. Under the Frensch et al. conception, if it were found that learning scores for a group trained under dual-task conditions were lower than those for a group trained under single-task conditions, this could be explained by suggesting that the single-task group possessed a degree of explicit knowledge about the sequence and that this was affected by the attentional

manipulation. Including a post-SRT explicit test in this study may help to address this issue.

To test the degree to which any sequence knowledge obtained was available to consciousness, participants completed two different free generation tests after the main SRT phase. Both used the same response keys as in the training and testing phase of the experiment. Participants completed an inclusion free generation test in which they were requested to try to generate the sequence they had responded to in the training phase of the experiment, and they completed an exclusion free generation test; here they were requested to generate any sequence other than the training sequence. If participants generated more triplets consistent with the training sequence under inclusion than exclusion instructions, then this may be taken as evidence of awareness of the training sequence, as participants could control their sequence knowledge (Destrebecqz & Cleeremans, 2001; Jacoby, Toth, & Yonelinas, 1993; Wilkinson & Shanks, 2004). If on the other hand the number of triplets generated was equivalent under inclusion and exclusion instructions, then this may be taken to demonstrate unconscious knowledge, as participants would in this case be unable intentionally to control the production of their sequence knowledge.

With the above methodological features, if implicit sequence learning proceeds independently of general attentional resources then learning scores for a group trained under single-task conditions should not differ statistically from learning scores of a group trained under dual-task conditions. If, on the other hand, implicit sequence learning is affected by the inclusion of a concurrent symbol-counting task, then it is expected that learning scores will be lower for the dual-task group compared with those of the single-task group.

### 2.1.1 Method

#### *2.1.1.1 Participants*

Thirty-eight volunteers (19 female, 19 male) from University College London took part in this study. Their mean age was 24.6 years with a range of 17-40. Participants were randomly assigned to one of two experimental groups, single or dual. They were paid £6 for their time upon completion of the experimental task; there was an additional award of a £15 book token for the six participants who performed best in the free-generation phase.

*2.1.1.2 Apparatus*

The experiment was fully automated and run on a PC using software written in Microsoft Visual Basic 6.0. Responses were made using a standard 'QWERTY' keyboard. Four boxes were arranged in a horizontal line just below the centre of the screen. Boxes were presented on a grey background; they were white in colour and were 12mm x 12mm in size. A 37mm gap separated each of the four boxes. On each target-location trial one of four possible text symbols ('x', '?', '$', and 'o') appeared in the centre of one of the boxes. The boxes will be referred to as locations 1-4 from left to right, respectively. Reaction times were measured using ExacTicks software in Visual Basic.

*2.1.1.3 Materials*

Two 12-item second-order conditional (SOC) sequences were used in this experiment; these are referred to as SOC1 (2-4-2-1-3-4-1-2-3-1-4-3) and SOC2 (3-4-3-1-2-4-1-3-2-1-4-2). For both sequences, predicting an item in the sequence depends on knowing the two prior target locations. The sequences were structurally identical. They were equated with respect to location frequencies (each location occurs 3 times), first order transition frequency (each location is preceded once by each of the other three locations), repetitions (no repetitions in either sequence), reversals (once only, i.e., 2-4-2 and 3-4-3). The only difference between the sequences is in their second-order conditional structure (e.g., 2-4 is always followed by 2 in SOC1 but by 1 in SOC2). The order of training sequences was counterbalanced so that half the participants saw SOC1 as the training sequence and half SOC2. During the training blocks, target location was specified by the assigned training sequence with a probability of .85 and by the alternative sequence with a probability of .15. For example if SOC1 was the training sequence, then the transition 3-1 was followed by a target at location 4 with a probability of .85 and by location 2 with a probability of .15. This procedure iterated on each trial and generated the next target based solely on the locations of the preceding 2 targets. A typical sequence might therefore be 3213413412423123 where the highlighted symbols refer to improbable locations (i.e., from SOC2) and the remainder to probable ones (from SOC1).

*2.1.1.4 Procedure*

Each participant was tested individually in a quiet room. Although all instructions were presented on screen, participants were informed that the experimenter would be outside the testing room and could be contacted if they had any queries regarding the experimental instructions. Participants in both groups were told that they were taking part in a simple reaction-time experiment that was designed to see how fast they could become at responding to a target. They were informed about the nature of the targets (the symbols x, o, ?, and $) and that they would appear in different location across the screen. On each trial, participants reacted to the location of a target by pressing keys V, B, N, and M for locations 1-4, respectively. Participants responded to locations 1 and 2 with the middle and index fingers of their left hands, respectively, and to locations 3 and 4 with the index and middle fingers of their right hands, respectively. Participants were instructed to respond to targets as quickly as possible and it was emphasised that they should avoid making any errors.

Firstly, all participants received 10 practice trials to become familiar with the task. The SRT display appeared with the letter symbols (V, B, N, and M) appearing under the corresponding box to aid correct finger placement. The 10 stimulus presentations showed each of the various symbols that participants had to respond to.

Each block consisted of 100 trials for a total of 4,000 trials. For each block the target-location trial began at a random point in the sequence. A target location trial ended when a participant made the correct key response, which resulted in the target being erased. Response latencies were measured from the onset of the target until the participant had made the correct response. On making the correct response for the target location the onset of the next target was immediate. If participants made an error this was signalled by a tone. The target remained until the correct response was made, and errors were recorded. At the end of each block dual-task participants were requested to type in their combined count of x's and ?'s and given feedback about their counting accuracy. If their accuracy was 100% they were commended and encouraged to maintain their high level of accuracy. If however it was less than 100% then they were asked to pay more attention to the counting task. Following completion of each block participants were told they could have a short pause and when ready they could initiate the next block by pressing the return key.

For all participants, the experimental procedure comprised a pre-training session on the SRT task under dual-task conditions, followed by 20 blocks of the SRT task in

the morning, and then 20 further blocks of the SRT task and a free-generation test in the afternoon. The interval between the morning and afternoon sessions was inserted to relieve fatigue. The purpose of including a pre-training period under dual-task conditions was to prepare participants for the dual-task procedure used in the training period. The crucial difference between the procedures for the two groups was the relative order and frequency of trial blocks performed under dual-task conditions (see Table 2.1).

Table 2.1. Experimental conditions for pre-training, each block of the SRT task, and post SRT. *S* single-task, *D* dual-task, *Inc* inclusion, *E* exclusion, numbers denote blocks. Participants had an extended break between blocks 20 and 21, which divided a morning and afternoon session.

| Experimental Group | Pre-Training | 1-9 | 10 | 11-19 | 20 | 21-29 | 30 | 31-38 | 39 | 40 | Post SRT |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Single-Task | D | S | D | S | D | S | D | S | S | D | Inc - Exc |
| Dual-Task | D | D | D | D | D | D | D | D | D | S | Inc - Exc |

Participants in the single-task group performed most blocks of the SRT task without performing a concurrent secondary task. The single-task group did, however, perform 4 test blocks under dual-task conditions (Blocks 10, 20, 30, and 40) so that learning effects could be investigated under identical conditions at regular intervals. Conversely, participants in the dual-task group performed the SRT task whilst performing the concurrent secondary task on every block except the final one (block 40). It is undesirable to subject the dual-task group to intermittent single-task test blocks as such a treatment may contaminate sequence knowledge acquired under conditions of divided attention by providing an opportunity for single-task learning.

The final test phase occurred at Blocks 39 and 40. For the single-task group block 39 was performed under single-task conditions and Block 40 was performed under dual-task conditions; for the dual-task group this order was reversed. Whenever a switch from single-task to dual-task conditions, or vice versa, was required, participants were instructed before the block of the necessary change. This procedure allows both single- and dual-task groups to be tested under both single- and dual-task conditions. Table 2.1 shows the experimental procedure in detail.

*2.1.1.5 Free Generation Test*

The free generation phase required participants to attempt to generate (inclusion) and to refrain from generating (exclusion) the training sequence. The order of these tests has been shown to have no effect (Wilkinson & Shanks, 2004), but in this experiment the order was counterbalanced across participants. Participants were informed that the targets followed a repeating sequence and were asked to press the keys 100 times, attempting to (i) freely generate the training sequence that they saw in the RT phase using the same keys, and (ii) freely generate a sequence that was as different as possible from the training sequence. They were told that each time they pressed a key, an "x" would appear in the appropriate box and that it would remain on the screen until a further key press was made, whence it would move to the next selected location immediately (i.e., the RSI was retained at 0 milliseconds). They were told not to worry if their memory for the sequence was poor, just to try to generate the sequence as best they could. The "x" moved to the corresponding location each time one of the four keys was pressed.

In the exclusion test, they still pressed the keys 100 times. They were instructed that if they could remember the sequence they should try to avoid generating it. In an effort to heighten motivation and encourage optimal performance, participants were told in advance that the top 6 performers on the free-generation test would each receive a £15 book token.

2.1.2 Results

*2.1.2.1 Response-Time Data for Blocks 1-40*

Figure 2.1 shows the mean RT data across all blocks of the SRT task for each condition, with probable and improbable target data plotted separately. The dramatic increases in RTs on Blocks 10, 20, 30 and 40 for the single-task group are due to the switch from a single-task to a dual-task procedure. The decrease in RTs on Block 40 for the dual-task group is due to the switch from a dual-task to a single-task procedure. Participants' mean RT data were entered into a global 3-way mixed-model ANOVA with the within-participants factors of Block (36 levels) and Target Probability (2 levels, probable and improbable), and the between-participants factor of Group (2 levels, single- and dual-task). Only 36 levels of Block were included in this analysis because the 4 dual-task test blocks (i.e. 10, 20, 30 and 39/40) were entered into a separate ANOVA (see below). The RTs for training blocks conducted under dual-task conditions

were in all cases considerably larger than those obtained under single-task conditions, $F(1, 36)=77.80$, with RTs decreasing with practice, $F(35, 1260)=58.36$. This practice effect appears to be more pronounced for the dual-task group, presumably because performance was initially hindered by the presence of the secondary task; accordingly, there was a significant Block × Group interaction, $F(35, 1260)=12.16$. From very early on, RTs for probable targets were lower than those for improbable targets, and this pattern was maintained across all training blocks in both groups, $F(1, 36)=168.01$, except for odd notable exceptions such as Blocks 2-4 of the dual-task group. There was no Probability × Group interaction, suggesting that during training, RTs to both probable and improbable targets were similar across groups. This consistent difference in RT (approximately 40 milliseconds on average) between probable and improbable targets is evidence that participants learned elements of the sequence and displayed priming effects for probable targets.

The three-way Block × Probability × Group interaction also reached significance, $F(35, 1260)=4.10$, indicating that the development of a probability effect across blocks is more robust in the single-task group.

*2.1.2.2 Difference Score Analyses*

The principal data of interest in the present study were the differences between RTs for probable and improbable targets on the test blocks. The primary test blocks were Blocks 39 and 40, but because blocks 10, 20 and 30 for the single-task group, and by default the dual-task group, were performed under dual task conditions they can provide additional data from which to analyse learning effects measured under identical conditions. The difference scores were calculated by subtracting the block mean probable RT from the block mean improbable RT for each participant and then averaging over all participants for that group. The difference scores for dual-task testing conditions are shown in Figure 2.2.

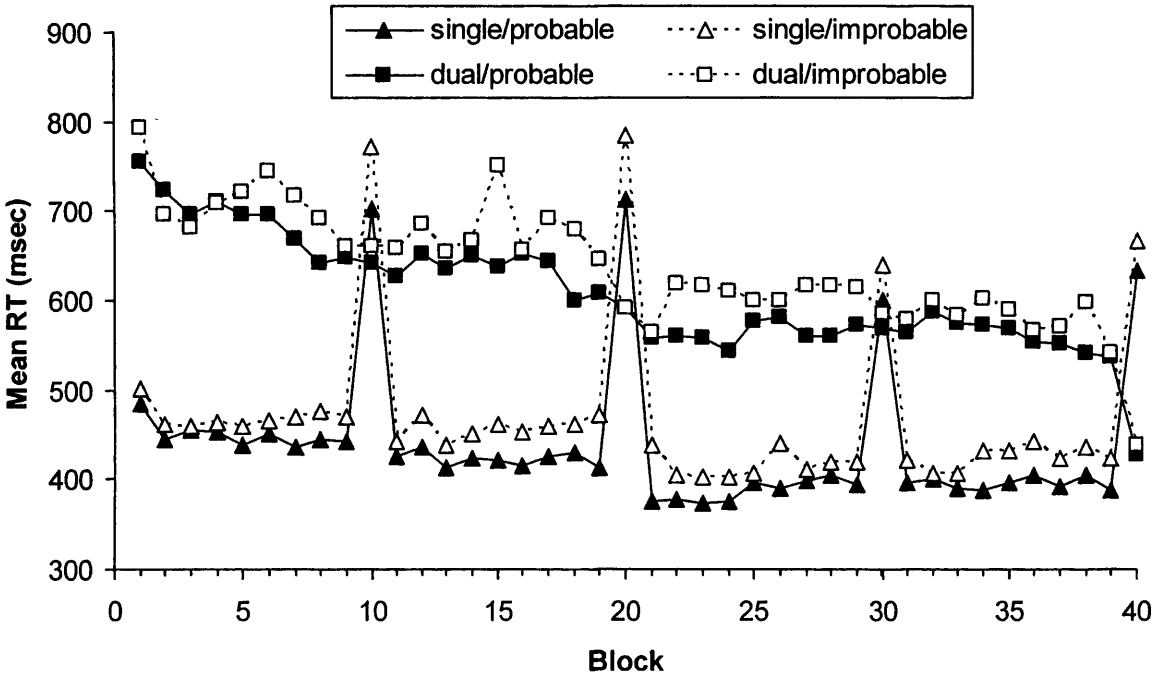


Figure 2.1 Mean RT for probable and improbable targets in Experiment 1 in a group (single) which performed the localization task alone on all blocks except blocks 10, 20, 30, and 40 and in a group (dual) which performed a concurrent symbol-counting task as well as the localization task across all blocks except block 40. The increase in RTs at blocks 10, 20, 30, and 40 for the single-task group represents the change from single- to dual-task conditions. The reduction in RTs on block 40 for the dual-task group represents the switch from dual- to single-task conditions.

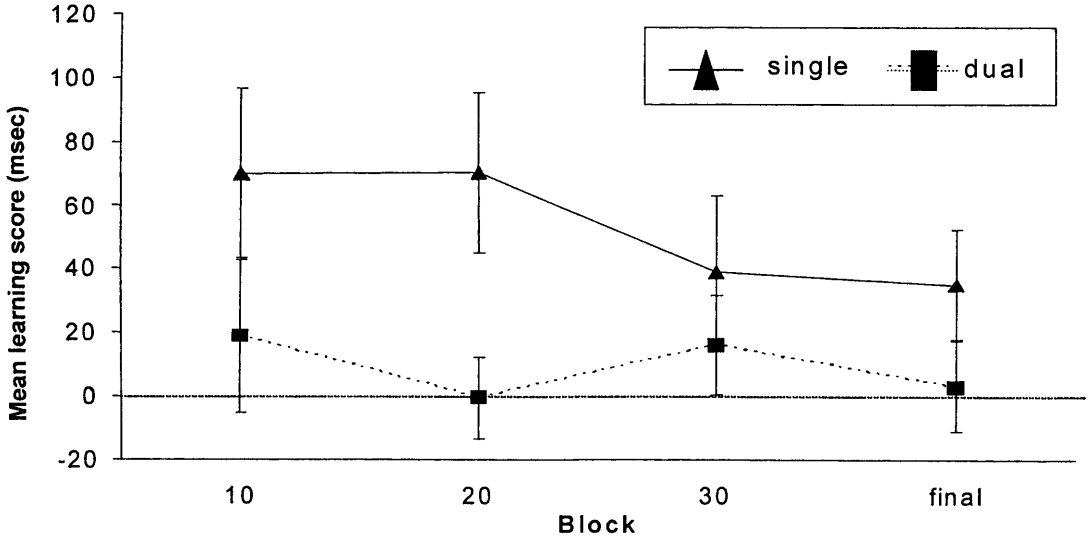


Figure 2.2 Mean difference between RTs for improbable and probable targets for both single- and dual-task training groups on dual-task testing blocks (Final refers to blocks 39 and 40 for dual- and single-task groups, respectively). Error bars depict the standard errors of the mean.

Consistent with the idea that sequence learning is impaired by a secondary task, difference scores for the single-task group were greater than for the dual-task group across all dual-task testing blocks. The final test block for the single-task group under dual-task testing was block 40, whereas for the dual-task group testing under dual-task conditions took place on block 39. With the exception of the dual-task group on block 20 all difference scores were greater than zero providing evidence of sequence learning.

The difference score data for the dual-task test blocks were entered into a 2 × 4 ANOVA with Group (single vs. dual) as a between-participants factor and Block (10 vs. 20 vs. 30 vs. Final Test) as a within-participants factor. The ANOVA results revealed a significant main effect for Group, $F(1, 36)=8.75$, suggesting that the two groups differed in the degree of sequence learning. The Block × Group interaction was not significant $F(3, 108)=1.49$, nor was there a main effect of Block, $F(3, 108)=1.53$.

*t* tests were computed between the difference scores of each group for each of the four test blocks. The probability values reported are one-tailed to conform to the prediction that learning scores would be greater for the single-task training group. There was no statistically significant difference between scores at Block 30, $p=.28$. However, there were reliable differences on Block 10, $t(36)=1.97$, $p<.05$, Block 20, $t(36)=3.42$, $p<.05$, and on the final test block, $t(36)=1.97$, $p<.05$, supporting the observation that mean difference scores were generally larger for the single-task training group when tested under dual-task conditions.

It is worth mentioning here the discrepancy between the relative degrees of single- and dual-task learning as revealed by the RT data during training, and that revealed by the learning scores on the test blocks. The absence of a Group × Probability interaction during training suggests that learning did not differ between groups, whereas the above analyses conducted on the learning scores revealed that single-task learning was greater than dual-task learning. There are several possibilities for this pattern of results. First, it is possible that a ceiling effect was prevalent in the single-task group, and thus participants simply could not perform the SRT task any faster. This would preclude the emergence of a maximal probability effect as revealed by RT differences between probable and improbable targets. This is supported to some extent by the observation that RT differences between probable and improbable targets are generally greater for the single-task group under dual-task testing conditions. This is an excellent example of why it is imperative to test both groups under identical conditions. However,

this also points to another possible explanation for the absence of a group learning effect: it is not desirable to compare RT differences obtained at one point on the RT scale, with those obtained at another. How learning is expressed at different performance speeds is not fully known, and may lead to incorrect conclusions being drawn. However, it is essential to acknowledge that the three-way interaction was significant, which suggests that there were differences in learning as revealed during training when the effect of practice was an additional factor.

Next it is important to analyse the difference scores for both groups under single-task test conditions. As can be seen from Figure 2.1, the difference between probable and improbable RTs for the single-task group on Block 39 was greater (35msec) than that for the dual-task group on Block 40 (11msec). Both these test blocks were conducted under single-task conditions. This difference was statistically reliable, $t(36)=2.11$, $p<.05$. Learning scores were therefore higher for the single- than for the dual-task training group when tested under single-task conditions.

*2.1.2.3 Error Data*

This was collected for response errors to probable and improbable targets on the SRT task, and for counting errors on the concurrent symbol-counting task. They are reported in Table 2.2. The symbol counting errors were calculated for each participant by computing the difference between their estimated count and the actual count on each block and converting into a percentage before averaging these percentages across all blocks. The errors reported below are means across all participants. For the symbol-counting task all participants were highly accurate on average with the range across all participants being 0-6% for the single-task group and 1-4% for the dual-task group. These error rates are well within the 10% error criterion adopted by Shanks and Channon (2002) and Cohen et al. (1990).

Table 2.2. Mean percentage errors for targets across all conditions, and in the symbol-counting task for dual-task blocks. Standard errors are in parentheses.

| | Mean Error Rates (%) | | |
|---|---|---|---|
| | Probable Targets | Improbable Targets | Counting Task |
| Task | | | |
| Single | 4.6(0.8) | 18.7(4.3) | 1.8(0.5) |
| Dual | 2.9(0.7) | 10.6(3.8) | 1.3(0.4) |

Target location error data were generally low for all conditions, with the single-task group producing the highest error rate overall for improbable target trials (18.7%). This latter finding is consistent with previous research (Schvaneveldt & Gomez, 1998) and suggests a greater degree of sequence knowledge leading to anticipation of probable target locations. However, it is also possible that because the single-task group were able to respond more quickly overall (because they were not hindered by the secondary task) they made more errors due to a speed-accuracy trade-off. The dual-task group on the other hand may actually have possessed the same degree of sequence knowledge but be less likely to respond as rapidly because they needed to first determine whether the target had to be counted.

*2.1.2.4 Free Generation Test Data*

To establish the extent to which sequence knowledge was consciously accessible, participants were requested to create sequences of 100 key presses under both inclusion and exclusion instructions. As Shanks and Johnstone (1999) have explained, there are numerous ways of analysing free generation data, but only one is adopted here.

To obtain inclusion and exclusion scores, each generated sequence was coded as 98 consecutive response triplets and the number of triplets that were consistent or inconsistent with the training sequence was calculated. For example if a participant was trained with SOC1 and generated the sequence 2-1-3-2 it was coded as two triplets: 2-1-3 and 1-3-2. 2-1-3 is consistent with respect to SOC1 but 1-3-2 is not. It is possible that some participants disregarded the free generation instructions in order to complete the experiment as quickly as possible (Wilkinson & Shanks, 2004). Such participants would have made perseverative key responses (e.g., pressing 1-2-3-4 repeatedly) throughout the tests; therefore all the data were scrutinised, and participants who showed such trends were removed from the analysis. The data of 5 participants were eliminated on these grounds. This left 17 participants in the single-task group and 16 in the dual-task group, though note that the statistical conclusions are identical when the entire sample is included.

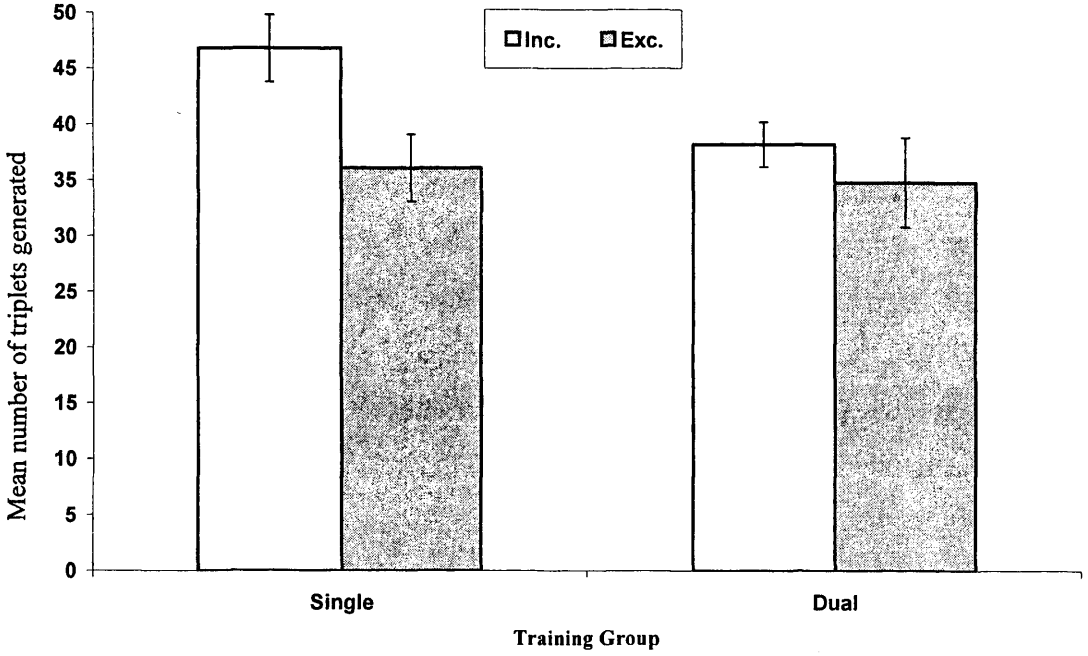


Figure 2.3 Mean number of second-order conditional (SOC) triplets generated that were consistent with the training sequence, under inclusion and exclusion instructions, for single-task and dual-task trained participants. *Error bars* depict the standard errors of the mean.

Figure 2.3 shows that the mean triplet generation score for the single-task training group is greater under inclusion than exclusion conditions. The same pattern, though less marked, is observed for the dual-task training group. These data were entered into a 2×2 mixed model ANOVA with Group (single- vs. dual-task training) as a between-participants factor and Test (inclusion vs. exclusion) as a within-participants factor. There was a statistically significant main effect of Test, $F(1, 31)=4.41$, but no main effect for Group, $F(1, 31)=1.05$; nor was there a statistically significant Group × Test interaction, $F<1$. These results indicate that to some extent participants possessed explicit knowledge of the probabilistic sequences used in this experiment. Thus, participants were able to intentionally control their sequence knowledge: they generated the training sequence to some extent when required, and could also refrain from generating it when requested. This suggests that both groups of participants had similar levels of intentional control over sequence knowledge.

To further explore exclusion performance, an additional analysis was conducted which followed the approach adopted by Destrebecqz and Cleeremans (2001), and was similar to that employed by Wilkinson and Shanks (2004). For each participant, the number of chunks generated for the training sequence was compared with the number of

chunks generated by the alternative sequence. As has been argued by the above authors, the rationale for analysing the data in this manner is that if sequence knowledge has been acquired implicitly during SRT learning, then it should be automatically produced during the generation test; that is, more chunks should be produced from the training sequence than from the alternative sequence, despite participants' attempts to refrain from generating the training sequence under exclusion conditions. An ANOVA on the exclusion free-generation data, with Group (single versus dual) as a between-participants factor, and Sequence (training versus alternative) as a within-participants factor, revealed no main effect of Group or Sequence and no interactions ($F$s < 1 in each case). To be certain that no differences exist between the number of chunks generated for the training and alternative sequences, $t$-tests were computed, which confirmed that there was no significant difference between the number of training (35.0) and alternative (33.8) chunks for the single group, $t(18) = 0.56$, and no difference between the number of training (36.4) and alternative (33.5) chunks for the dual group, $t(18) = 1.11$. These results support the conclusion that both groups had control over sequence knowledge acquired during the SRT task, as they do not (implicitly) produce more triplets consistent with the training sequence than those consistent with the alternative sequence. This result is consistent with the findings of Wilkinson and Shanks (2004), who obtained evidence that probabilistic sequence learning under no-RSI conditions is to some extent under intentional control. The present results extend those found previously by demonstrating that SRT learning under dual-task conditions also leads to sequence knowledge that is under conscious control to approximately the same extent as that obtained under single-task training.

### *2.1.3 Relationship Between Sequence Learning and Free Generation*

Together, the RT and free generation data of this experiment are somewhat paradoxical; the RT data reveal a difference in learning between dual- and single-task conditions, whereas both groups showed similar levels of conscious sequence knowledge according to the inclusion-exclusion difference in the free generation tasks. What is the significance of this apparent dissociation?

The most plausible conclusion is that the magnitude of the inclusion-exclusion difference is somewhat insensitive to variations in the extent of sequence knowledge as indexed by RT difference scores. However, it is important to acknowledge another possibility, namely that the data reflect distinct learning mechanisms. Perhaps the group

data result from the combination of distinct patterns from two sub-groups of participants: (1) a subgroup that has explicit knowledge (i.e., a positive inclusion-exclusion difference), and in which sequence learning is attention-demanding (i.e., an effect of the secondary task), and (2) a subgroup that has no explicit knowledge (i.e., no inclusion-exclusion difference), and in which sequence learning is not attention-demanding (i.e., no effect of the secondary task). If that is the case, then one would expect to see a much smaller difference between single- and dual-task RT difference scores in participants who showed the smallest inclusion-exclusion difference. In fact the data do not show this: a median split was conducted on each of the two groups on the basis of the magnitude of their inclusion-exclusion difference, and then the mean RT difference score averaged across blocks 10, 20, 30, and Final for each subgroup was calculated, which revealed that these scores are 67 and 13 msec for the single- and dual-task groups, respectively, for the subgroup with the largest inclusion-exclusion difference and 46 and 8 msec for the single- and dual-task groups, respectively, for the subgroup with the smallest inclusion-exclusion difference. An ANOVA revealed a reliable effect of training conditions (single/dual), $F(1, 29)=7.50$, $p=.01$, but no effect of inclusion-exclusion difference (high/low) and no interaction, $F<1$ in each case. Notably, the difference in the sub-group with poor free generation performance is reliable, $t(16)=1.92$, $p<.05$ (one-tailed). Hence the effect of the secondary task is unrelated to the degree of conscious knowledge as indexed by the free generation test.

It is important to emphasize that this dissociation between the two measures is not in itself evidence for distinct learning systems. As has been shown elsewhere (Shanks, 2005; Shanks & Perruchet, 2002; Shanks et al., 2003), a single-system model in which a common underlying variable determines the magnitude of both sequence knowledge and conscious knowledge is capable of predicting a complete lack of correlation between them if there is independent noise or error in the processes which translate that variable into each type of response.

Finally, another analytic method for determining attentional effects on implicit and explicit learning would involve estimating the separate contributions of these based on the assumptions of the process dissociation procedure (Jacoby et al., 1993). This method has not been adopted here because of doubts over the validity of these assumptions (Ratcliff, McKoon, & Van Zandt, 1995; Tunney & Shanks, 2003).

### *2.1.4 Discussion*

This study has investigated the effects of attentional load on implicit sequence learning. The RT data suggest that learning under dual-task conditions detrimentally affects sequence learning. This is supported by the finding that RTs to probable and improbable targets differed more for single-task participants than for dual-task ones. The effect occurred when sequence knowledge was tested both under single- and dual-task conditions. Therefore it is inferred that the single-task participants demonstrated greater sequence learning.

The lower level of learning found for dual-task participants compared to single-task participants is consistent with the results of Shanks and Channon (2002), in that reduced attention had a detrimental effect on sequence learning. In that study, however, deterministic rather than noisy sequences were used and it is widely agreed that the latter are preferable to ensure that the involvement of explicit processes is minimised (Cleeremans & Jiménez, 1998). Like the results of Shanks and Channon, the present finding that learning is greater for single-task participants even when tested under dual-task conditions casts doubt on some earlier studies (e.g., Curran & Keele, 1993, Experiment 2) which reported equivalent learning effects in single- and dual-task groups tested in this way. Overall, the findings conflict with the notion that sequence learning is able to proceed without the use of attentional resources (Cleeremans & Jiménez, 1998; Frensch, 1998; Frensch et al., 1998). Moreover, the symbol counting task provides a purer effect of reduced attention on sequence learning as it is suggested to be free from some of the disruption effects associated with tone counting (Jiménez & Méndez, 1999; Stadler, 1995). Thus the conclusion drawn here is at variance with previous claims that implicit learning may proceed independently of processing capacity, yet is consistent with the proposal that implicit learning does rely on attentional resources (Hsiao & Reber, 1998; Perruchet & Vinter, 2002) and with research that has failed to find evidence for the automatic nature of implicit memory (Mulligan & Hornstein, 2000).

The free generation data revealed that both dual- and single-task groups had a degree of intentional control over their sequence knowledge: both were able to refrain from generating information from the training sequence when asked to do so during the exclusion test. Both groups of participants were able to generate more sequence-consistent triplets under inclusion than exclusion instructions. The data from the process dissociation comparison of the free generation task challenges the traditional view that

implicit learning yields an unconscious knowledge base. As indicated by the differences in triplet generation under inclusion and exclusion instructions, both groups' sequence knowledge was, to some degree, conscious. Fuller discussion of the inclusion/exclusion comparison and its interpretation can be found in Destrebecqz and Cleeremans (2001) and Wilkinson and Shanks (2004). This finding is consistent with results obtained by Shanks, Wilkinson, & Channon (2003) who also reported that training on probabilistic sequences allows the development of explicit sequence knowledge.

The RT results conflict with the findings of Jiménez and Méndez (1999) who reported no effect of divided attention on sequence learning. They employed a similar secondary task to that used in this study along with probabilistic sequences. Jiménez and Méndez used a finite-state grammar to generate their sequences (Cleeremans & McClelland, 1991), thus producing far noisier sequences than those in the present experiments. Why are there contrasting results between this study and that of Jiménez and Méndez (1999)? Several possibilities are considered below.

One possibility is due to differences in the way that the degree of learning was measured. The experiment reported here tested participants under both dual- and single-task conditions, regardless of training conditions. Jiménez and Méndez did not adopt this procedure. With the exception of their Experiment 2, learning was measured simply by analysing mean RT differences between probable and improbable targets. They found that there was no statistically significant interaction between training conditions and target type; that is, RT differences between probable and improbable targets did not differ as a function of training conditions. At the end of training on Jiménez and Méndez's Experiment 2 the secondary task was removed so that the dual-task group could be tested under single-task conditions. Again, there was no statistically reliable difference in learning between groups. The above methods for measuring learning are unsatisfactory because, in the first case, there is no reason to assume that learning scores at lower RT values (single-task group) relate in any meaningful way with learning scores at higher RT values (the dual-task group). This is more specifically known as a scaling effect. In the present experiment this problem was circumvented by testing both groups under dual-task conditions. Single-task testing alone is not adequate because of the likelihood of a ceiling effect: RTs to probable targets simply cannot be faster than our visuo-motor system will allow, and thus the full degree of learning may not be expressed. This supposition is supported in the present experiment as learning scores for the single-task training group were generally higher under dual-task conditions. This

finding further undermines the notion of Frensch's (1998) 'suppression hypothesis', and supports Shanks, Wilkinson and Channon's (2003) finding that dual-task procedures do not reduce the expression of sequence learning.

Another potential source of the discrepancy between the present results and those of Jiménez and Méndez, is that, as explained in the introduction to this chapter, training in the Jiménez and Méndez study was very long. RTs in their experiments were averaged over 3,100 trials (compared with only 100 trials in the present study) which is close to the total amount of training given in here. The present treatment thus allows for a more sensitive analysis of learning. It is clear from Figure 2.1 that both groups show a robust learning effect by Block 10, which amounts to just 1,000 trials. Jiménez and Méndez's method does not provide a detailed insight into the rate of learning, and the corollary of this is that the analysis is likely to be insensitive to any differences during early stages of learning when the secondary task is probably most demanding (i.e., before it may become automatised or before task-sharing is possible) and thus when the disparity in the rate of acquisition is most prominent.

A third reason is that the present study, unlike Jiménez and Méndez's, did not arrange a predictive relationship between shapes and targets. This was discussed in detail in the introduction to the present chapter, and thus will not be dealt with here.

Although the present data provide convincing grounds for doubting the attentional-independence claim, there should be acknowledged yet further possibilities: The first is that the present results came about not because the secondary-task is attention demanding, but because it disrupts the temporal structure of the task; perhaps the symbol-counting processes engaged in by participants disrupt their perception of the organisation of the primary sequence (Hsiao & Reber, 2001; Rah, et al., 2000; Stadler, 1995). The problem with this suggestion is that it is hard to imagine how the attentional-independence claim can be tested without introducing a secondary task that disrupts the timing and organisation of the location sequence. Proponents of the claim would need to suggest some way of testing it that avoids this eventuality. Yet in contrast to this claim, the experiments in Chapter 3 of this thesis (see also Rowland & Shanks, 2006a) do not report any affect on learning from a task that does alter the perception of the organisation of the sequence. Along very similar lines, an additional argument might suggest that the counting process impairs learning because the act of counting itself is disrupting, in that participants need to switch between two tasks, one visuo-motor (the SRT task) and the other involving some sort of articulatory rehearsal and updating in

working memory (the counting task). Perhaps this is disrupting not because it consumes resources (which it almost certainly does), but because of the switching between concurrently competing goals. The challenge is to develop an SRT task that involves an additional task-switching task that loads minimally on attentional resources, whilst still not introducing additional stimuli into the procedure (recall that the symbol-counting task was designed to avoid this kind of interference). Perhaps if participants had to classify the symbols and state their decision out loud, yet not count them, this may potentially impose a suitable task-switching demand, without a working memory demand?

Alternatively, a second possibility is that perhaps there are two separate forms of learning, a fast variety (revealed in the present experiment, and those in Shanks et al., 2005) which makes attentional demands, and a slower variety (revealed in Jiménez and Méndez' experiments) and which does not. This possibility is consistent with the data presented here and needs to be considered in further research.

In conclusion this study has challenged both the notion that implicit sequence learning is able to proceed without attentional resources and the idea that it is accurately characterised by a lack of awareness of the products of learning.

# Chapter 3

## Sequence Learning and Selection Difficulty

The present chapter investigates implicit learning by adopting a novel approach that considers some recent advances in the study of attention. As reviewed in Chapter 1, it has been common to conduct experiments on the role of resource attention and selective attention in implicit learning. However, it is widely recognised that attention is more appropriately viewed as a collection of phenomena (Allport, 1992; Driver, 2001), and concordantly, in an attempt to account for the spectrum of attentional variations, researchers have proposed models of attention that are composed of several mechanisms (e.g., Duncan, 1999; Posner & Boies, 1971; Posner & Petersen, 1990; Wickens, 1984). Potentially, these may differ in their relevance to implicit learning, and thus, further clarifying the way in which attention interacts with implicit learning is the focus of the experiments presented in the present chapter (Experiments 3-5 also are reported in Rowland & Shanks, 2006a).

Sequence learning is a perceptual-motor task that requires the successful operation of both visual selective attention to the locations (and sometimes the identity) of stimuli, and selective control over responses made to the target stimuli. In this view, the SRT task comprises two stages: visual selection of the target, and response selection of the correct motor output. If there are separate attentional mechanisms devoted to these aspects of the task then perhaps their effects on sequence learning are dissociable? Of particular interest then is the notion of two capacity-limited mechanisms of attention (Johnston et al., 1995; Lavie et al., 2004), one responsible for selective aspects of perception (input attention), and the other responsible for control processes, such as the allocation of attention, maintenance of processing priorities, and response selection (central or resource attention). This mechanism of attentional control is thought to depend crucially on working memory (Baddeley, 1996; de Fockert et al., 2001). Such componential models of attention have rarely been considered in implicit learning research, yet such views would undoubtedly provide a richer framework for studying the effects of attention on sequence learning.

An important question therefore is whether implicit sequence learning is sensitive to load on some attentional mechanisms and not others. The study described in Chapter 2 demonstrates that load on central attention impairs sequence learning, and research indicates that selective attention too is necessary for implicit learning in the

SRT task to occur (Jiménez & Méndez, 1999). However, in these studies, resources and selection have been treated as independent, when in fact, as contemporary models of attention have postulated, they are inextricably linked. For example, experiments that led to the perceptual load theory proposed by Lavie and her colleagues (e.g., Lavie, 1995; Lavie & Tsal, 1994) showed that selection becomes increasingly focussed as attentional capacity is consumed (i.e., when the goal-relevant perceptual task becomes more demanding). By considering the perceptual load theory of attention, it becomes obvious that questions such as whether learning requires either central resources or selective attention may need to be reconsidered. In the context of implicit learning, it may be more useful instead to ask whether loading a central attentional mechanism (such as a control mechanism like working memory) or loading a perceptual mechanism (such as a sensory input mechanism involved in the selection and filtering of visual information) might have differing effects on the development of sequence learning. If it is the case, as Jiménez and Méndez (1999) have suggested, that selective attention to targets in the SRT task is necessary for implicit learning about them to occur, then it follows that the selection process is a fundamentally important event in this kind of learning. A potentially interesting line of enquiry therefore is to ask what are the effects of disrupting these vital selection operations. Specifically, does the selection of targets and the filtering of distractors in the SRT task compete for the same pool of resources that are vital for optimal implicit sequence learning?

In the field of implicit *memory* there has been considerable effort to elucidate the effects of several attentional manipulations (see Mulligan & Brown, 2003, for a review of this literature), and the current evidence appears to suggest that the availability of both input attention and central attention are vital if implicit memory is to be intact. However, some evidence using the contextual cueing paradigm supports a view of implicit *learning* that operates independently of input attention: Jiang and Chun (2001) found a robust contextual cueing effect even when participants had to ignore irrelevant distractors in the display. However, Jiang and Chun did not set out to investigate this process specifically - for instance, they did not have a control group who performed the task in the absence of distractors - and correspondingly, conclusions about the role of input attention in implicit learning can only be drawn cautiously from their study. Clearly, detailed considerations of attentional processes in implicit learning research are important for advancing our understanding, and thus the experiments in the present chapter used an attention-demanding task designed to load input attention, but not load

SRT task to occur (Jiménez & Méndez, 1999). However, in these studies, resources and selection have been treated as independent, when in fact, as contemporary models of attention have postulated, they are inextricably linked. For example, experiments that led to the perceptual load theory proposed by Lavie and her colleagues (e.g., Lavie, 1995; Lavie & Tsal, 1994) showed that selection becomes increasingly focussed as attentional capacity is consumed (i.e., when the goal-relevant perceptual task becomes more demanding). By considering the perceptual load theory of attention, it becomes obvious that questions such as whether learning requires either central resources or selective attention may need to be reconsidered. In the context of implicit learning, it may be more useful instead to ask whether loading a central attentional mechanism (such as a control mechanism like working memory) or loading a perceptual mechanism (such as a sensory input mechanism involved in the selection and filtering of visual information) might have differing effects on the development of sequence learning. If it is the case, as Jiménez and Méndez (1999) have suggested, that selective attention to targets in the SRT task is necessary for implicit learning about them to occur, then it follows that the selection process is a fundamentally important event in this kind of learning. A potentially interesting line of enquiry therefore is to ask what are the effects of disrupting these vital selection operations. Specifically, does the selection of targets and the filtering of distractors in the SRT task compete for the same pool of resources that are vital for optimal implicit sequence learning?

In the field of implicit *memory* there has been considerable effort to elucidate the effects of several attentional manipulations (see Mulligan & Brown, 2003, for a review of this literature), and the current evidence appears to suggest that the availability of both input attention and central attention are vital if implicit memory is to be intact. However, some evidence using the contextual cueing paradigm supports a view of implicit *learning* that operates independently of input attention: Jiang and Chun (2001) found a robust contextual cueing effect even when participants had to ignore irrelevant distractors in the display. However, Jiang and Chun did not set out to investigate this process specifically - for instance, they did not have a control group who performed the task in the absence of distractors - and correspondingly, conclusions about the role of input attention in implicit learning can only be drawn cautiously from their study. Clearly, detailed considerations of attentional processes in implicit learning research are important for advancing our understanding, and thus the experiments in the present chapter used an attention-demanding task designed to load input attention, but not load

working memory, in order to examine the effect on sequence learning of disrupting the selection process.

## 3.1 Overview of the Experiments

To date, SRT studies have used simple displays whereby selecting a target stimulus has been undemanding (but see Cock et al., 2002). This is clearly unlike the noisy natural environment with which an implicit learning system is familiar (and adapted to) and precludes any detailed investigation of the selection process. Accordingly, the selection process in the present experiments was made more complex by the inclusion of distracting stimuli. In all the experiments (except Experiment 5) the target stimulus was a red circle that appeared on each trial at one of four screen locations, the order of which was determined by a probabilistic sequence. Participants simply had to respond to the target location as quickly as possible by pressing a key assigned via a compatible mapping. The target was immediately erased and appeared instantaneously at the next prescribed location. Participants performing under single-stimulus conditions had only the target present during the SRT task. In contrast, participants performing under dual-stimulus conditions also had distractor stimuli (e.g., a green square) appear on each trial at one or more of the same four possible screen locations as could the target. These distractor items were irrelevant to the task and were to be ignored.

The inclusion of a salient irrelevant distractor on each trial requires that attention be narrowed down to the target to prepare for a response. The spatial uncertainty before stimulus onset encourages attention to be spread over the four possible locations, but upon appearance of the stimuli attention must focus on the target while irrelevant objects are filtered. The delay in response times under such conditions has been termed the 'filtering cost' (Kahneman, Treisman, & Burkell, 1983; Treisman, Kahneman, & Burkell, 1983). The effect of a filtering cost on implicit sequence learning has not previously been explored.

The purpose of the experiments was to investigate what effects, if any, such distractor items have on primary sequence learning. In all four experiments participants were trained under single (i.e., no distractors) or dual/triple (i.e., with distractors) conditions, and then sequence learning was measured under single conditions (in Experiment 2 sequence learning was also measured under dual conditions). The logic behind the experiments is as follows. The inclusion of irrelevant distractor stimuli

increases the difficulty of target selection and thus imposes a filtering cost. The filtering cost is the increase in response time to the target (compared with conditions where no distractor is present) and is due to the spreading of attention over multiple stimuli (i.e., target and distractors), and the subsequent narrowing of the focus of attention on the target. This task thus loads attention in a different way from the traditional methods of concurrent working memory tasks: here, no response to, or selection of, the distractor is necessary. It is postulated, in line with the componential models presented above, that the presence of irrelevant distractors loads a sensory input mechanism of attention. It is worth noting here the similarity between the inclusion of distractor stimuli in the SRT task and the notion of perceptual load. In one sense it is accurate to characterise the manipulation as increasing perceptual load, but this general usage is avoided here because the term has often been applied to conditions where selection difficulty is increased by the addition of task-*relevant* items (e.g., Lavie & Tsal, 1994), whereas the present experiments manipulated selection difficulty by the inclusion of *irrelevant* items.

It should be mentioned that some authors have argued that sequence learning is primarily a *motor* task. If this were so, then including distractors at the perceptual level may seem an inappropriate manipulation for studying the effects of input attention on implicit learning. However, a wealth of evidence shows SRT learning to be largely *non-motoric*. Transfer tests, for example, show that learning acquired in the SRT task can transfer from three fingers to one finger (Cohen et al., 1990), from fingers to the ipsilateral arm, and across modalities from manual to vocal responding (Keele, Jennings, Jones, Caulton, & Cohen, 1995). Grafton, Hazeltine, and Ivry (2002) showed transfer from the non-dominant to the dominant hand in the SRT task, and previously the same authors showed transfer from the fingers to the arm (Grafton et al., 1998). Willingham, Wells, Farrell, and Stemwedel (2000) used a crossed-hands transfer test and showed that learning transferred across effectors (i.e., it was not tied to a sequence of finger movements) but did not transfer to a new sequence of keys. Willingham and colleagues concluded that what was learned in the SRT task is a sequence of response locations. Moreover, Cock et al. (2002), Howard, Mutter, and Howard (1992), Mayr (1996), and Remillard (2003) have shown that learning can occur in the SRT task in the absence of motor responses (i.e., with observation alone). The evidence therefore suggests that perceptual processes play a large part in SRT learning, and that visual distractors are thus suitable for investigating the selection processes. Their presence

may, for example, interfere with the essential process of associating contiguous sequence-response locations.

Both single and dual/triple learners performed the same primary task (i.e., responding to target locations in the SRT task). One question the present experiments sought to answer therefore is whether selective attention to the target stimulus is sufficient for optimal sequence learning to occur (Jiménez & Méndez, 1999). If this were the case, then learning under single- and dual-stimulus conditions should be equivalent. Alternatively, the imposition of a filtering cost may be detrimental to sequence learning, and thus learning scores would be higher for single than for dual/triple groups.

## 3.2 Experiment 2

In this experiment participants were trained for 13 blocks on the probabilistic SRT task under either single (no distractors) or dual (one distractor present) conditions. Then for a further 2 blocks the single group performed under dual conditions and the dual group performed under single conditions. This design allowed comparisons of sequence learning between the single and dual learners to be made under both single and dual testing conditions. This procedure has previously been adopted (Shanks et al., 2005) because it avoids the problem that the expression of sequence learning may be suppressed when participants are trained and tested under demanding (i.e., dual-task) conditions (Frensch et al., 1998, 1999).

If interference from an irrelevant distractor item attenuates learning then learning should be greater for the single group compared with the dual group.

### 3.2.1 Method

#### *3.2.1.1 Participants*

Forty-four individuals, the majority of whom were University College London undergraduates (30 women and 14 men, mean age = 21.7 years, range 19-29) volunteered to take part in this experiment. They were assigned randomly to one of two experimental groups, single ($n$ = 22) or dual ($n$ = 22). They were each paid £3 for completing the task. For all the experiments reported in this chapter, all participants volunteered, were from the same population, and were paid £3.

*3.2.1.2 Apparatus*

The hardware and software for the experiments reported in this chapter was identical to that of Experiment 1, with the exception of the stimulus locations. Instead of four boxes, the stimulus locations were replaced by four horizontal black lines (length = 8 mm, breadth = 1mm).

*3.2.1.3 Materials*

These were identical to those of Experiment 1, with two exceptions. First, 4 SOC sequences were used: SOC1 (2-1-2-3-4-1-3-2-4-3-1-4), SOC2 (4-1-4-3-2-1-3-4-2-3-1-2), SOC3 (3-1-3-2-4-1-2-3-4-2-1-4), and SOC4 (4-1-4-2-3-1-2-4-3-2-1-3). SOC1/SOC2 and SOC3/SOC4 formed the two sequence pairs for the purposes of generating training and alternative sequence pairs. Second, in the SRT task the target stimulus was a red circle, 3 mm in diameter, and the distractor was a green square, with sides of length 5mm.

*3.2.1.4 Design and Procedure*

The experiment employed a 2 × 2 design, where the between-participants factor was the 2 levels of training conditions (single or dual) and the within-participants factor was the 2 levels of target probability, probable or improbable. Assignment to single or dual conditions was counterbalanced.

Participants performed 15 blocks of 100 trials each. For the single group, the first 13 blocks of the task were performed with only the target stimulus present, whereas for the dual groups, the first 13 blocks were performed with both the target stimulus and the distractor present. During the final two blocks (14 & 15), the single group switched to performing the SRT task under dual conditions, whereas the dual groups switched to performing under single conditions.

Participants were informed that the experiment was a simple choice reaction-time study investigating response performance to a repetitively presented stimulus. Participants were instructed to respond to the targets as quickly as possible whilst attempting to keep errors to a minimum. They were not told that the sequence of locations was structured. Participants were required to place the middle and index fingers of their left hand on the V and B keys of the keyboard for responses to locations 1 and 2, respectively, and the index and middle fingers of their right hand on the N and

M keys for responses to locations 3 and 4, respectively. They then completed 25 practice trials of the single-stimulus SRT task. Following this, participants in the dual groups were informed that an additional stimulus, a green square, would appear on each trial during the SRT task, and that they were required as best as possible to ignore it. Participants then pressed a key to begin the experimental task.

The target stimulus followed a sequence of locations as prescribed by a probabilistic SOC sequence (see Materials above). One of the pairs of SOC sequences was randomly assigned to each participant, and the sequence determining the probable transitions was randomly selected from the pair, with the other sequence of the pair determining improbable transitions. For the dual group the distractor stimulus also followed a probabilistic sequence, determined by the SOC pair not designated as the training and alternative sequences for that participant. On those trials where the target and the distractor appeared at the same location, the red circle occupied the foreground, placed centrally within the green square so that the edges of the square were clearly visible.

At the end of block 13, instructions appeared on the screen informing participants to continue performing the task as they had been doing, but that a change would occur: the single group were informed that an additional stimulus, a green square, was to be added and that they should attempt to ignore it; the dual group were informed that the green square was to be removed.

During the SRT task, each block began with a target-location trial selected from a random point in the training sequence. For dual blocks, the initial distractor location was also selected randomly. Henceforth, locations followed their respective sequences.

Each trial, and reaction-time measurement, began at onset of the target, and ended when the participant made a correct location response which triggered immediate offset of the stimuli, and the commencement of the next trial (the response-stimulus interval was 0 milliseconds because it appears that this is likely to engage implicit learning processes more fully; Destrebecqz & Cleeremans, 2001). If participants pressed the incorrect key, the target stayed at the current location until a correct response was made. Participants were permitted to have a short break at the end of each block and when ready initiated the next block by pressing the Return key.

Avoidance of target localisation errors was emphasised in all the experiments reported in this chapter. Participants were provided with auditory feedback (a beep)

each time an error occurred, and were reminded before the test blocks to try as hard as possible to avoid making errors.

All other details of the procedure were identical to those of Experiment 1 in Chapter 2.

### 3.2.2 Results

#### *3.2.2.1 Response-Time Data for Training Blocks (1-13)*

For all the experiments reported, RTs on error trials were excluded from the analyses and RT data were collapsed across all training sequences. The alpha level is set at .05, two-tailed, throughout, unless otherwise stated.

Figure 3.1 presents the mean RTs obtained for single and dual groups for both probable and improbable trials across the training blocks (1-13). During training, dual participants generally responded more slowly to targets than did single participants although this difference was negligible by the end of training. This shows that the distractor did have a negative impact on response speed but that its effect was not sustained over the duration of training. An analysis of variance (ANOVA) on the data from blocks 1-13 with Block, Group, and Target Probability as factors revealed main effects of Block, $F(10, 420) = 4.53$, $MSE = 2{,}235$, and Target Probability, $F(1, 42) = 56.64$, $MSE = 1{,}515$, and statistically significant interactions of Block × Group, $F(10, 420) = 4.27$, $MSE = 2{,}235$, and Block × Target Probability, $F(10, 420) = 2.80$, $MSE = 686$. As a whole, these results indicate that RTs decreased more across block in the dual group, and that, indicative of probabilistic sequence learning, disparity in RTs for probable and improbable targets (with probable RTs being reliably lower) increased over training for both single and dual groups. However, neither the Target Probability × Group interaction, nor the Block × Target Probability × Group interaction reached significance (both $F$s < 1), which indicates that sequence learning did not differ between groups over training.

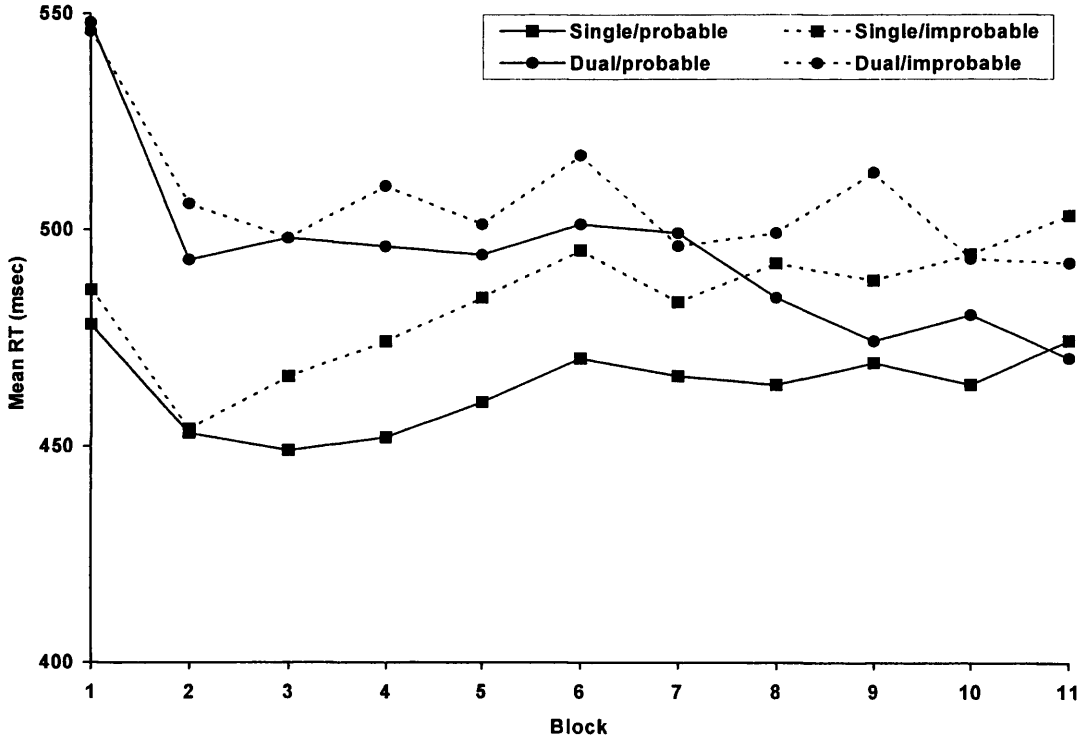


Figure 3.1 Mean reaction time (RT) for probable and improbable targets across the training blocks of the SRT task in Experiment 2 for the single- and dual-task groups.

*3.2.2.2 Learning Scores*

Figure 3.2 presents the learning score data for both single and dual learning groups each tested under single and dual conditions. These scores are a measure of the difference (in milliseconds) between probable and improbable trials, collapsed across both test blocks for each test type. Blocks 12 and 13 were designated test blocks for the single group under single testing conditions and for the dual group under dual testing conditions. Blocks 14 and 15 were designated test blocks for the single group under dual conditions and the dual group under single testing conditions.

It is clear from Figure 3.2 that difference scores are much lower for dual testing blocks than for single testing blocks for both single and dual groups. This was confirmed by an ANOVA with Group and Testing Conditions as factors revealing a main effect of Testing Conditions, $F(1, 42) = 6.75$, $MSE = 1{,}636$. This latter result presumably reflects a suppression of the expression of learning when performing the SRT task under dual conditions. More interestingly, there was no effect of Group, and no Group × Testing Conditions interaction ($F$s<1); thus, regardless of testing conditions, there was no difference in learning scores between single and dual groups. Both the single group, $t(21) = 3.43$, $p < .01$, and the dual group, $t(21) = 4.96$, $p < .01$, obtained

learning scores under single testing conditions that were greater than zero. Under dual testing conditions only the dual group obtained a learning score greater than zero, $t(21) = 3.02$, $p < .01$. Presumably the single group had difficulty expressing sequence knowledge under dual testing that was acquired during single training, and consequently the learning scores for this group did not differ significantly from zero, $t(21) = 0.77$, *ns*.

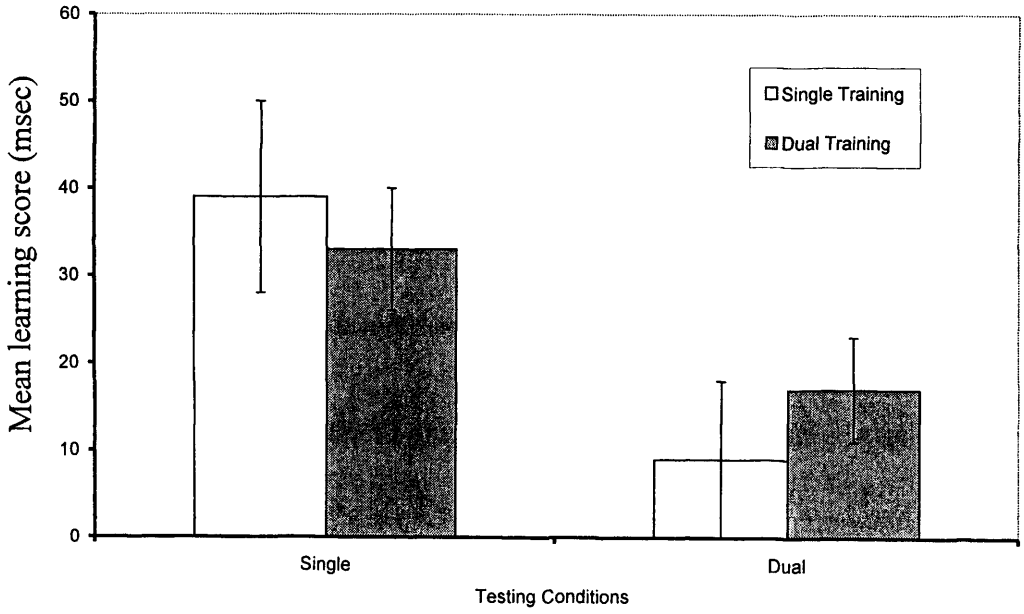


Figure 3.2 Mean learning score for single- and dual-task training groups under single- and dual-task testing conditions. *Error bars* depict standard errors of the means.

*3.2.2.3 Error Data*

Error rates (i.e., incorrect target localization responses) were higher for improbable target trials than for probable trials for all groups, although generally low (i.e., means across all blocks were: probable = 10.13%, improbable = 14.50%). This is characteristic of probabilistic sequence learning, and presumably reflects the anticipatory responses to sequence locations that may accompany learning (Schvaneveldt & Gomez, 1998; Shanks et al., 2003, 2005). There were no statistically reliable group differences in error rates regardless of test conditions ($F<1$). The pattern of the error data reported here was the same in the other experiments in this chapter, and will not be discussed further.

### 3.2.3 Discussion

The results of this experiment demonstrate that when tested under identical testing conditions (i.e., single and dual), the degree of sequence learning is the same regardless of whether participants were trained with or without the presence of distractors. Crucially, because RTs were overall slower under dual conditions than under single conditions, it is reasonable to assume that the irrelevant stimulus was an effective distractor. Hence it appears that sequence learning is not impaired by the imposition of a filtering cost.

There are several weaknesses to this experiment. First, testing conditions under single and dual conditions occurred at different stages of the SRT task for the different groups, and it is possible that this created an advantage for the dual group by allowing extended learning before being tested under single conditions. Given that the single group had a numerically greater learning score under single testing than did the dual group, this difference might be greater if tested after identical periods of training.

Second, it is clear from Figure 3.1 that by block 8 the dual-stimulus group were responding at approximately the same speed as that of the single group. Indeed, filtering costs are known to diminish with practice (Kahneman et al., 1983), hence training should be shortened so that learning can be measured when the distractor is having greater impact. Third, both target and distractor appeared at the same location on a considerable number of trials, but evidence shows that filtering costs are eliminated with spatial overlaying (Kahneman et al., 1983; Treisman et al., 1983). This can be avoided by using a distractor sequence that prevents the distractor from appearing at the same location as the target on any trial. These limitations are addressed in the following experiment.

## 3.3 Experiment 3

The aim of this experiment was twofold: first, to re-examine the effect of a single distractor on sequence learning; and second, to examine whether the nature of the distractor *sequence* affects learning of the primary sequence. This latter issue is of some importance in this study because current evidence raises the possibility that random and structured distractor sequences may have differential effects on sequence learning. In the case of random stimuli, they appear to disrupt learning by altering the temporal and structural organization of the primary sequence (Rah et al., 2000; Jiménez & Vázquez,

2005), whereas the presence of systematic stimuli may alleviate the effects of distraction simply by being predictable and thus more easily ignored or integrated into the primary sequence (Schmidtke & Heuer, 1997). Thus, it is necessary to determine whether the method of distractor presentation has effects that complicate the primary goal. At face value, these effects are unlikely to be significant, but it is essential to reject this possibility from the outset as a potential problem. In contrast to the above studies, the stimuli in the present experiments appeared simultaneously with the targets (instead of during the inter-stimulus interval as has often been the case), required no punctate responses, and were the same on every trial. Moreover, the systematic stimuli followed a different probabilistic sequence from that of the target, and so despite being highly irregular in their structure, the possibility of the sequences being correlated, and thus learning of the primary sequence being enhanced or augmented (Batsell, 2000) is minimised. Under these conditions it is highly unlikely that the temporal or structural organisation of the primary sequence will be disrupted, or that the systematic stimuli will be too easily ignored. Nevertheless, two different dual groups were included, with distractors that either followed a systematic sequence or were random.

To address the problem raised in the discussion of Experiment 2, namely that the distractor ceases to be effective after extended training, participants were trained for only 8 blocks on the probabilistic SRT task under either single or dual conditions. Then for a further 2 blocks the single group continued with the same task, whereas the dual groups performed under single conditions. There was no dual testing in this experiment for two reasons: first, so that single testing could occur at the same stage of the SRT task for both groups; and second, because there appeared to be a suppression effect in Experiment 2 that disadvantaged the single group.

It is also important to determine whether negative priming arises during these experiments. In this context, negative priming would result if target responses were slower on trials where the current target location had been occupied by a distractor on the previous trial, compared with those where the current target location had been unoccupied on the previous trial. One explanation for this effect is that distractor stimuli are substantially perceived (Lavie & Fox, 2000; Rees et al., 1997), and that active inhibition is required to moderate interference from them (Cock et al., 2002; Houghton & Tipper, 1994; Tipper & Cranston, 1985). Evidence of negative priming would therefore indicate that the distractors are receiving attention, both for processing and for preventing interference.

### 3.3.1 Method

#### *3.3.1.1 Participants*

Thirty-six individuals (22 women and 14 men, mean age = 23.2 years, range = 18-35) took part in this study. They were assigned randomly to one of three experimental groups, single ($n$ = 12), dual-random ($n$ = 12), or dual-structured ($n$ = 12).

#### *3.3.1.2 Apparatus and Materials*

These were identical to Experiment 2.

#### *3.3.1.3 Design and Procedure*

The experiment employed a 3 × 2 design, where the between-participants factor was the 3 levels of training conditions (single, dual-random or dual-structured) and the within-participants factor was the 2 levels of target probability, probable or improbable. Assignment to single or dual conditions was counterbalanced.

Participants performed 10 blocks of 100 trials each. For the single group, the first 8 blocks of the task were performed with only the target stimulus present, whereas for the dual groups, the first 8 blocks were performed with both the target stimulus and the distractor present. During the final two blocks (Blocks 9 & 10), the single group continued performing the SRT task as normal, whereas the dual groups switched to performing under single conditions.

For the dual-structured group the distractor stimulus followed a probabilistic sequence, determined by the SOC pair not designated as the training and alternative sequences for that participant. On those trials where the target and the distractor appeared at the same location, the red circle occupied the foreground, placed centrally within the green square so that the edges of the square were clearly visible. For the dual-random group the distractor stimulus appeared at locations pseudorandomly with the constraint that it could not occupy the same location as the target (to avoid reduced filtering costs by spatial overlaying, Kahneman et al., 1983; Treisman et al., 1983) and could not appear at the same location successively.

At the end of Block 8 for the dual group, instructions appeared on the screen informing participants to continue performing the task as they had been doing, but that the distractor would be removed.

In all other respects, the procedure was identical to Experiment 2.

### 3.3.2 Results

#### *3.3.2.1 Response-Time Data for Training Blocks (1-8)*

Figure 3.3 presents the mean RTs obtained for all groups for both probable and improbable trials over all blocks of the SRT task. During training, dual participants responded more slowly to targets than did single participants, thus showing that the distractor stimulus did have a negative impact on response speed.

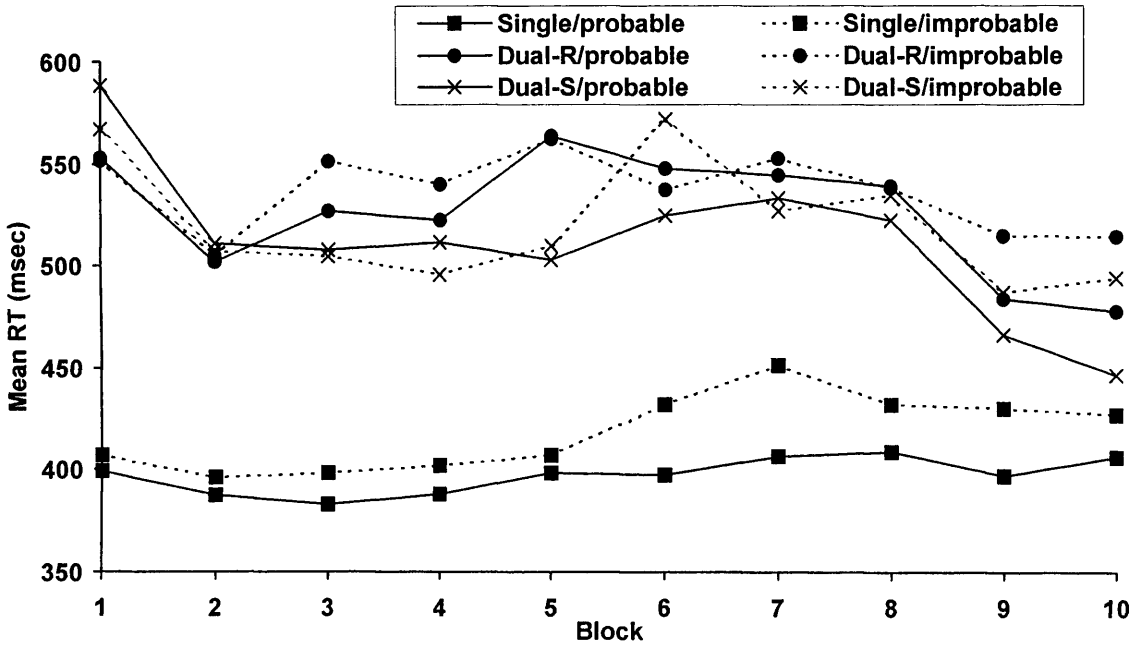


Figure 3.3 Mean RT for probable and improbable targets across all blocks of the SRT task in Experiment 3 for the single, dual-random, and dual-structured groups. Blocks 1-8 were the training blocks and blocks 9 and 10 were the testing blocks.

The single group showed a stable probability effect, whereas for the dual groups this is somewhat noisier. Learning tends to be quite rapid in the probabilistic SRT task (Schvaneveldt & Gomez, 1998; Shanks et al., 2005), generally being evident by Block 3 and not increasing much in magnitude thereafter, so the critical statistical evidence for learning is a main effect of target probability rather than a Probability × Block interaction. Overall, there was a main effect of target probability, $F(1, 33) = 5.15$, $MSE = 2{,}191$, and no Probability × Group or Probability × Block × Group interactions ($Fs < 1.98$), confirming that all groups learned the sequence to approximately the same degree. The Probability × Block interaction was nonsignificant, $F(7, 231) = 1.43$, *MSE*

= 969. Nevertheless, the probability effect in the single-task group was significant by the final blocks (see below) and was not significant at Block 1, $t(11) = 1.70$ (the same pattern occurs in all subsequent experiments in this chapter), confirming the development of sequence learning. RTs were much slower for the dual groups compared with the single group because of distraction: main effect of group, $F(2, 33) = 17.5$, $MSE = 59{,}542$, $p < .01$. Both dual groups had approximately equal RTs during training, suggesting that both were distracted to a similar degree.

*3.3.2.2 Learning Scores*

For all the remaining experiments in this chapter the mean learning scores were calculated by subtracting the probable RT from the improbable RT for each participant on each of the final 2 blocks (under single conditions), collapsing across blocks, and averaging over all participants. They are shown in Figure 3.4. Both dual groups had identical learning scores (*M*s = 34 milliseconds), and thus there does not appear to be an effect of the underlying nature of the distractor sequence on SRT learning. The single group had a slightly lower learning score than the dual groups ($M = 27$ milliseconds). Hence sequence learning does not seem to be disrupted by a selection cost. All three groups generated learning scores greater than zero: single, $t(11) = 3.37$, $p < .01$; dual-structured, $t(11) = 2.70$; and dual-random, $t(11) = 2.56$.

*3.3.2.3 Negative Priming*

This was investigated by comparing the mean RT for target trials where the location of the target was preceded by the presence of a distractor with target trials where the location was not preceded by the appearance of a distractor. It was anticipated that negatively primed (NP) trials would be slower than non-NP trials, because participants would inhibit responses to that location on the immediately previous trial.

NP and non-NP trials were calculated for the dual-group across blocks 1-8. For NP trials, RTs were averaged over all trials across all 8 blocks where the target appeared in a location where a distractor had appeared in the same location on the immediately previous trial. The same procedure was adopted for non-NP trials, except that RTs were averaged for trials where the target appeared at a location that was unoccupied by a distractor on the immediately previous trial.

Consistent with this notion and the idea that distractors did consume attention, NP trials were slower ($M = 541$ milliseconds, $SE = 7.32$) than non-NP trials ($M = 520$

milliseconds, $SE$ = 8.89), $t(22)$ = 6.58, $p < .01$. Both dual-stimulus groups were collapsed into a single group for this analysis.

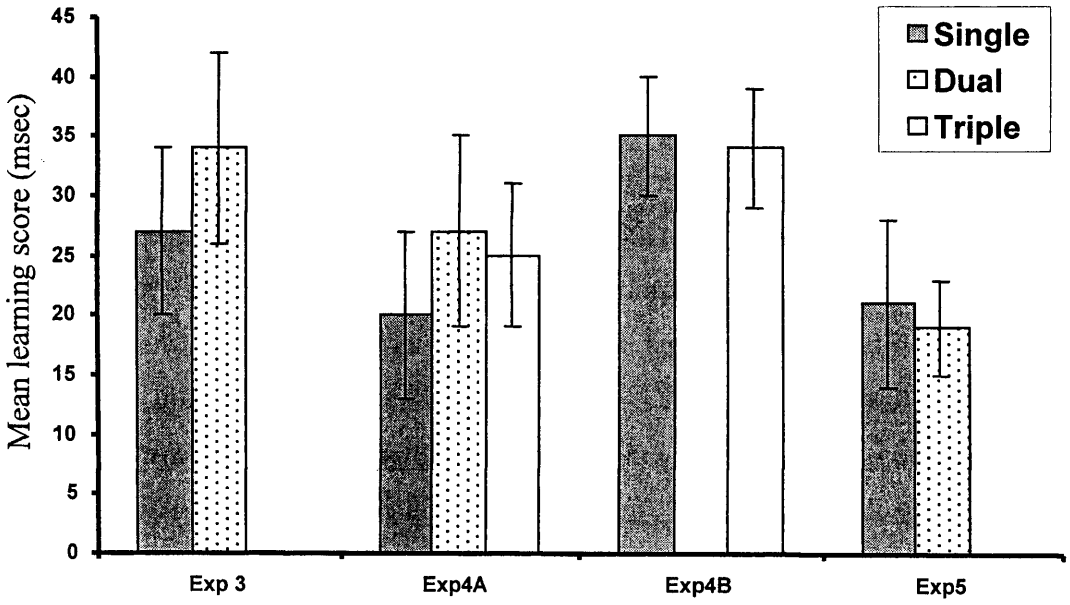


Figure 3.4 Mean difference between RTs for probable and improbable targets under single-stimulus testing conditions for all groups in Experiments 3, 4A, 4B and 5. Error bars depict the standard error of the mean. Learning scores were identical for both dual-stimulus groups from Experiment 3 (i.e., dual-random and dual-structured) and thus collapsed into a single dual-stimulus group in the graph.

### 3.3.3 Discussion

The results of this experiment demonstrate that when compared under identical single testing conditions, the degree of sequence learning is the same regardless of whether participants were trained under either single or dual conditions. Crucially, because RTs were overall slower under dual conditions, and because there was strong evidence of negative priming, it is apparent that the irrelevant stimulus was an effective distractor. Hence, the finding from Experiment 2 that sequence learning is not impaired by the imposition of a filtering cost has been replicated.

Moreover, there is no evidence that the nature of the distractor sequence has an effect on primary sequence learning, which allows us to draw conclusions about the

effects of selection difficulty without concern for this issue. It was discussed above why the results were likely to differ from other research into the effects of distractor sequences, and thus no further consideration will be given to this matter here. For the remaining experiments distractor sequences are random. This choice was made because there is evidence that distractor sequences can be learned (Cock et al., 2002) and it is not clear to what extent this may compete for attentional resources with the primary task, especially over extended training durations (as in Experiment 4B).

## 3.4 Experiments 4A and 4B

These experiments attempted to replicate and extend the findings of the previous two experiments. Both Experiments 2 and 3 may underestimate the effects of selection difficulty either because (a) training was too long and hence the effect of the filtering cost on learning diminished with practice, or (b) the presence of one highly salient distractor did not load input attention sufficiently. Experiment 4A therefore sought to increase the impact of the distractors on sequence learning. First, training was reduced to only 6 blocks (any less and it is unlikely that sufficient learning in the single group would have developed – see Figure 3.3), and second, 2 levels of selection difficulty were employed. The low level was identical to that of Experiments 2 and 3 (a green square). The high level was created by presenting two irrelevant stimuli: A green circle and a red square. Note that each of these stimuli shares one property with the target stimulus (a red circle), yet none with each other. A conjunction search was thus required to identify the target stimulus.

Another possible reason for a lack of an effect in Experiment 3 is that training was not long enough, and thus the single group did not learn the sequence to a high enough degree to detect a difference between dual and single learning. Experiment 4B integrates both a conjunction search for the target and greater practice. This manipulation ensures that high selection difficulty is maintained for the dual group whilst allowing for greater development of sequence learning in the single group.

Together these experiments enable an investigation of various manipulations of selection difficulty on sequence learning: a single distractor with little practice; a conjunction search for the target with little practice; and a conjunction search with greater practice.

### 3.4.1 Method: Experiments 4A and 4B

#### *3.4.1.1 Participants*

A further 80 individuals (48 women and 32 men, mean age = 24.3 years, range = 18-41) took part in this study. The first 56 participants were assigned randomly to one of three experimental groups for Experiment 4A, single-stimulus ($n$ = 19), dual-stimulus ($n$ = 19), and triple-stimulus ($n$ = 18). The remaining 24 participants were assigned randomly to either the single-stimulus ($n$ = 12) or triple-stimulus ($n$ = 12) groups for Experiment 4B.

#### *3.4.1.2 Apparatus and Materials*

These were identical to Experiment 3 except that the triple groups were presented with two distractor stimuli: a filled red square (sides = 5 mm) and a filled green circle (3 mm in diameter).

#### *3.4.1.3 Design and Procedure*

The procedure was identical to Experiment 3, except where detailed below. Experiment 4A adopted a between-participants design with three conditions. The groups were single-stimulus, dual-stimulus and triple-stimulus. Participants performed the SRT task for 8 blocks. For the first 6 blocks of the task, the dual-stimulus group was presented with an accompanying distractor stimulus (a green square) and the triple-stimulus group was presented with two accompanying distractor stimuli (a red square and a green circle). These were removed for the final two blocks. Similar to the dual-random group of Experiment 3, the distractor stimuli did not follow a sequence, instead appearing at locations pseudorandomly.

Experiment 4B was identical to Experiment 4A except that only 2 groups were used (single and triple) and the SRT task was performed for 14 blocks (12 training blocks, 2 testing blocks).

### 3.4.2 Results for Experiment 4A

#### *3.4.2.1 Response-Time Data for Training Blocks (1-6)*

Figure 3.5 shows the mean RT data across the training blocks (1-6) of the SRT task for all groups.

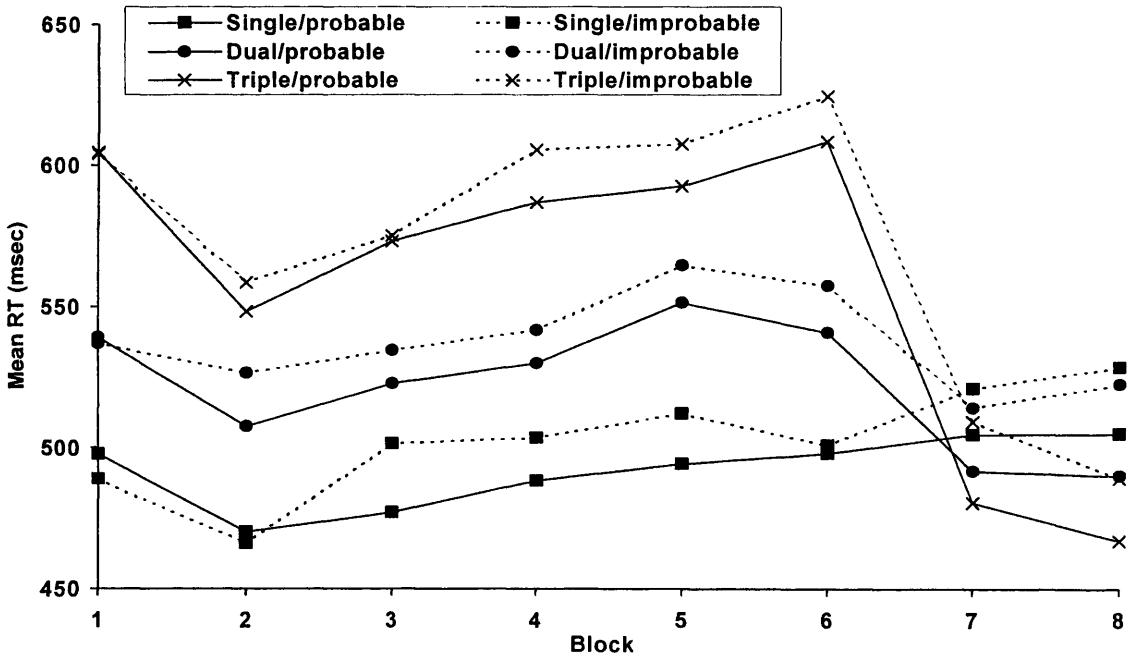


Figure 3.5 Mean RT for probable and improbable targets across all blocks of the SRT task in Experiment 4A for the single-, dual-, and triple-stimulus groups. Blocks 1-6 were the training blocks and blocks 7 and 8 were the testing blocks.

Participants' mean RTs were entered into an ANOVA with the factors Block (6 levels), Target Probability (2 levels), and Group (3 levels, single/dual/triple). There was a main effect of Group, $F(2, 53) = 4.84$, $MSE = 113{,}092$, which it is clear from Figure 3.5 reflects the fact that RTs increased (approx. 80 milliseconds per distractor) across groups as the number of distractors increased. Most importantly, there was a highly significant main effect of Target Probability, $F(1, 53) = 23.67$, $MSE = 695$, $p < .01$, yet no Target × Group interaction, which suggests that all groups learned the sequence, and that the expression of this learning was identical under all levels of distraction. No other interactions reached, or approached, significance ($F$s < 1.86).

*3.4.2.2 Learning Scores*

These were calculated as in the previous experiments, and are shown in Figure 3.4. The data were entered into a one-way ANOVA, which revealed no statistically significant difference ($F$<1). This confirms that learning scores were similar across groups. All three groups obtained learning scores greater than zero: single, $t(18) = 2.50$, dual, $t(18) = 2.81$, and triple, $t(17) = 3.32$, $p < .01$.

### 3.4.3 Results for Experiment 4B

#### *3.4.3.1 Response-Time Data for Training Blocks (1-12)*

Figure 3.6 shows the mean RT data across all blocks of the SRT task, for both groups. The analyses on the training blocks (1-12) reveal no group differences in learning (i.e., a lack of a Target Probability × Group interaction) and a highly significant main effect of Target, $F(1, 22) = 20.33$, $MSE = 3{,}189$, $p < .01$, confirming that a conjunction search over extended practice greatly impairs response times to the target.

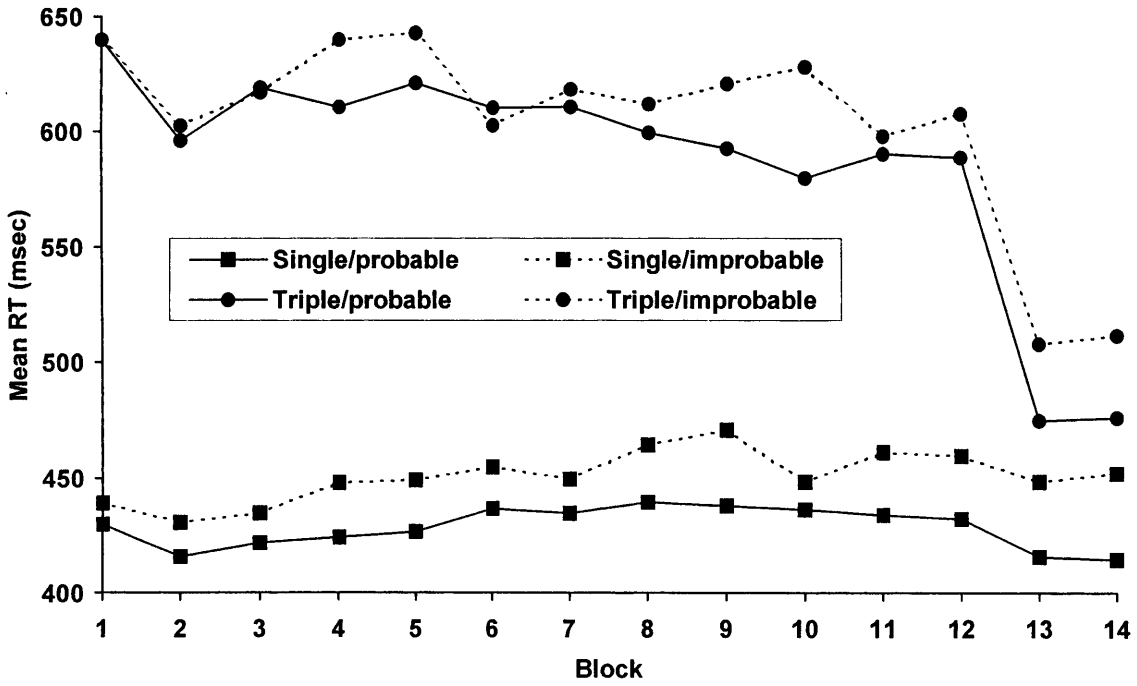


Figure 3.6 Mean RT for probable and improbable targets across all blocks of the SRT task in Experiment 4B for the single- and triple-stimulus groups. Blocks 1-12 were the training blocks and blocks 13 and 14 were the testing blocks.

#### *3.4.3.2 Learning Scores*

The single-stimulus group generated learning scores greater than zero, $t(11) = 3.44$, $p < .01$, as did the triple-stimulus group, $t(11) = 4.74$, $p < .01$, and these were virtually identical (35 milliseconds vs. 34 milliseconds; see Figure 3.4; $t(22) = 0.04$, *n.s.*).

#### *3.4.3.3 Negative Priming*

There was a large difference between the means (37 milliseconds), $t(11) = 7.35$, $p < .01$, with NP trials being slower ($M = 628$ milliseconds, $SE = 6.28$) than non-NP

trials ($M$ = 591 milliseconds, $SE$ = 10.91). All 12 participants in the triple-stimulus group showed strong evidence of negative priming. Note that the data for computing NP (locations of the distractor on each trial) were not recorded for Experiment 4A.

### 3.4.3 Discussion of Experiments 4A and 4B

The results of Experiment 4A show that the presence of either one or two distractors during the SRT task does not attenuate sequence learning. This indicates that increases in selection difficulty do not interfere with SRT learning, even when participants are forced to select the target on the basis of a conjunction of features. This finding is replicated in Experiment 4B over longer training duration. The strong evidence of negative priming in Experiment 4B is further confirmation that distractor items were receiving attention (Cock et al., 2002; Lavie & Fox, 2000). It is worth considering the possibility that two distractors that do not share features with either the target or with each other (e.g., if the target were a red circle and the distractors comprised a green square and a blue triangle) may increase the selection demand above that of a conjunction search. However, this is unlikely given that the detection of feature conjunctions appears to require greater attentional involvement than does the detection of single features (e.g., Treisman & Gelade, 1980) and thus this potential confound was not addressed experimentally.

## 3.5 Experiment 5

It is possible that the demands on input attention are not sufficiently high even during a conjunction search. Perhaps the target stimulus is too easily identified amongst the distractors, and after a few hundred trials places negligible load on input resources. Indeed, Johnston and Dark (1986) noted that not all selection processes are equal in the demands they place on attention. For this reason a new approach to disrupting the selection process is adopted here. In this experiment, the targets and distractors are members of two separate categories, either vowels or consonants. The stimulus presentation is random within categories, so on each trial the form of the target and distractor is unpredictable, requiring selection of the target to be based on category membership. Selecting the target under these conditions is likely to be more effortful than performing a conjunction search, and thus this experiment builds on the previous experiments by examining the effects of further increasing selection demands.

### 3.5.1 Method

#### *3.5.1.1 Participants*

A further 24 individuals (13 women and 11 men, mean age 24.2, range = 18-35) took part in this study. They were assigned to group single ($n$ = 12) or dual ($n$ = 12).

#### *3.5.1.2 Apparatus and Materials*

The target items were vowels presented in upper case, and the distractors were consonants chosen for their relative perceptual similarity (i.e., shape and features) to a vowel with which it was paired: A/H; E/F; I/J; O/Q; U/V. The letters were size 14 Arial font.

#### *3.5.1.3 Design and Procedure*

There were two groups, single and dual. The single group responded to the location of a vowel. The dual group responded to the location of a vowel and ignored a consonant distractor. On each target trial one of the five vowels was chosen randomly. There were 8 blocks of the SRT task. All other features of the experiment were consistent with those already reported.

### 3.5.2 Results

#### *3.5.2.1 Response-Time Data for Training Blocks (1-6)*

Figure 3.7 shows the mean RT data across all blocks of the SRT task for both groups. An ANOVA on the data from blocks 1-6 with Block, Group, and Target Probability as factors revealed a main effect of group, $F(1, 22) = 15.55$, $MSE = 144{,}059$, $p < .01$. Plainly, the difference in RTs (approximately 150 milliseconds) between the single- and dual-stimulus groups shows that the presence of the irrelevant consonant was very disrupting. There was also a Block × Group interaction, $F(5, 110) = 4.26$, $MSE = 1{,}123$, which reflects the fact that RTs for the dual-stimulus group improved over training as participants became increasingly able to ignore the distractor (or, put another way, select the target). This likely reflects learning to respond to the shape of the vowels without having to processes the stimulus category. The implications of this are briefly discussed below. Surprisingly there was no main effect of target probability, but there was a Target Probability × Group interaction, $F(1, 22) = 12.86$, $MSE = 3{,}203$, $p < .01$. This pattern suggests that the single group learned the sequence during training whereas

the dual group did not (as is clear from Figure 3.7). In an analysis restricted to RTs on Block 6, there was a main effect of Probability, $F(1, 22) = 10.15$, $MSE = 402$, $p < .01$, which also interacted with Group, $F(1, 22) = 6.16$. Thus, bearing in mind that the difference score data below show clearly that learning took place, the dual-stimulus group likely suffered greater disruption because of the difficult nature of the category-selection task.

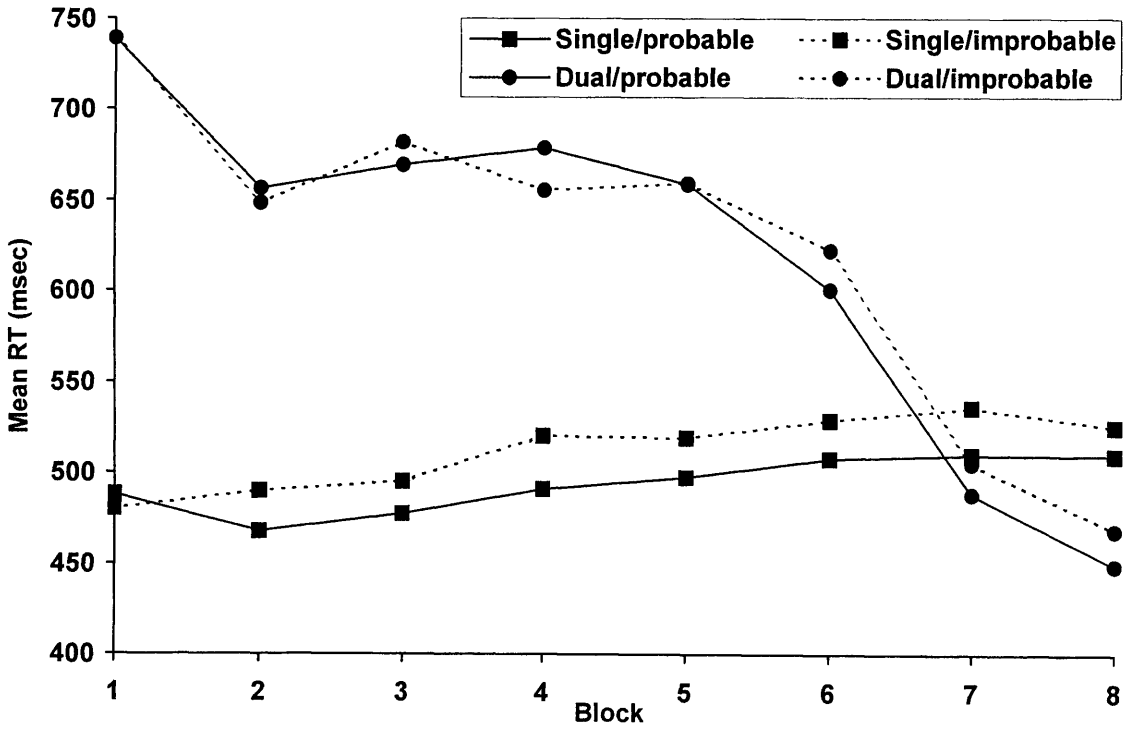


Figure 3.7 Mean RT for probable and improbable targets across all blocks of the SRT task in Experiment 5 for the single- and dual-stimulus groups. Blocks 1-6 were the training blocks and blocks 7 and 8 were the testing blocks.

*3.5.2.2 Learning Scores*

These are shown in Figure 3.4. Consistent with the results from the other experiments in this chapter, disruption to the selection process did not impair sequence learning: difference scores were 21 milliseconds for the single-stimulus group and 19 milliseconds for the dual-stimulus group, $t(22) = 0.03$, *ns*. Both single- and dual-stimulus groups generated learning scores greater than zero, $t(11) = 2.67$, $p < .05$, $t(11) = 3.12$, $p < .01$, respectively.

### 3.5.3 Discussion

Experiment 5 demonstrates that even when more effortful selection requiring category identification is used with a greater filtering cost (the dual-stimulus cost was greater than that seen in the earlier dual-distractor groups), there is still no detrimental effect on primary sequence learning. In this experiment, unlike the earlier ones, there was some evidence of an expression deficit when distractors were present: Evidence of learning was weaker in the dual group in training than at test, consistent with the possibility that learning was "released" at test by removal of the distractors. However, it is possible that the category-selection task became easier after repeated exposure to the vowels and that target selection may have been achieved by attending to the form of the stimuli instead of their category membership. If this is correct, then the experiment does not address fully the effects of category-selection on SRT learning. However, this is not problematic for the current study for two reasons. First, the RT data are testament to the increased demands (compared with Experiments 4A and 4B) induced by the category-selection task. Second, a purely category-based selection task would undoubtedly impinge on working memory and thus involve the very attentional processes that the present experiments aimed to avoid.

## 3.6 Power Analyses for Experiments 2-5

Despite large performance effects between groups, these experiments have revealed no differences between single-, dual-, or triple-stimulus learning scores. As always when interpreting null results, it is necessary to determine the power of the study to detect an effect of selection cost on sequence learning, should one exist. In Experiments 3 and 4A, where there were two experimental groups, power was computed between the single-stimulus group and the collapsed experimental groups. Thus, all power analyses were computed for a two-group (control and experimental) between-participants comparison. For Experiments 2, 3, 4A, 4B, and 5, power to detect a medium effect was .49, .32, .54, .40, .32, respectively. Power to detect a large effect was .83, .60, .88, .72, .60, respectively. For all analyses, alpha was .05, $n$ was as reported in the *Method* sections, hypotheses were one-tailed, and the measure of effect size was Cohen's $d$ (Cohen, 1988). For a medium sized effect, combined power over all four experiments reached .93. This was calculated by multiplying the probability of committing a Type II error (ß) across all experiments (= .13), and then subtracting from 1. Combined power for a large effect was .99 and for a small effect was .54. To gain an

estimate of the size of effect that might be expected when investigating the effects of distraction on SRT learning, the effect size was calculated between single- and dual-task SRT learning from the results of Experiment 1 in Shanks et al. (2005) in which a secondary symbol-counting task was added. A Cohen's *d* of .78 was obtained, which is a medium to large effect. Apart from the nature of the attentional manipulation, Shanks et al.'s experiment used the same sequences, similar sample sizes, and comparable training lengths (1,000 trials) to those used here.

Although the above calculations suggest that the conclusion drawn from the test stage data of these experiments is unlikely to reflect a Type II error, it is important also to bear in mind that in none of the experiments reported here (except for the borderline case of Experiment 5) was there any evidence of greater learning in the single-stimulus groups during the training stage. Despite the fact that the presence of distractors profoundly slowed down RTs, they did not attenuate the development of faster responses to probable versus improbable targets.

Interestingly, this finding suggests that the presence of distractors not only did not affect sequence learning, but they did not affect the expression of that learning either. Evidence that the expression of sequence learning is not affected by the presence of performance-impairing preparations has been reported before (Shanks and Channon, 2002), but in previous instances there has been evidence of an accompanying detriment to learning, which the present experiments do not reveal. This implies that probabilistic sequence learning can be optimal even under conditions where the structure of the sequence is altered, and is testament to the robustness of such learning.

## 3.7 Discussion of Experiments 2-5

The experiments in the present chapter investigated the effects of increasing the difficulty of target selection on probabilistic sequence learning. Selection difficulty was manipulated by adding irrelevant distractor stimuli to the SRT task that were to be ignored. Filtering costs arise when irrelevant items in a display delay responses to target items by vying for attention, despite the target items being highly salient (Kahneman et al., 1983; Treisman et al., 1983). Across five experiments no evidence was found of impaired learning for groups trained on the SRT task in the presence of distractor stimuli even though filtering costs were apparent. The data show that under identical single-stimulus testing conditions, learning (as indexed by RT differences between probable and improbable targets) was identical for groups trained with and without

distractors. This result was obtained under several variations: low selection difficulty with just a single salient distractor (Experiments 2, 3 and 4A); high selection difficulty with two distractors each sharing one feature with the target, hence target selection being based on a conjunction of features (Experiments 4A and 4B); and, finally, a more effortful selection requiring category discrimination (Experment 5). The evidence strongly suggests that implicit sequence learning is not affected by disruption to the selection process.

At first glance, the findings of the present chapter appear to conflict with the result in Chapter 2 that sequence learning is attenuated under conditions of load or distraction (see also Shanks et al., 2005). In both the latter and the present experiments, the same sequences and training durations were used, and similar overall performance decrements (in terms of elevated RTs) for the dual-task groups were obtained. In contrast, a major difference between the procedures is that Shanks et al.'s experiments and the one reported in Chapter 2 involved a secondary working memory task designed to load central resources, whereas the present experiments used a distracting task with a minimal working memory component, instead designed to load input attention. This suggests that implicit sequence learning is selectively impaired by attentional load, being sensitive to load on central resources, but relatively unaffected by load on input attention. Hence, implicit sequence learning is resistant to disruption of the selection process, yet highly dependent upon the availability of working memory. This pattern provides corroborating evidence for the componential view of attention described in the Introduction to the present chapter, and extends the findings of Jiang and Chun (2001) by comparing groups trained with and without distraction, by systematically increasing the degree of distraction, and by investigating a different implicit learning task.

These results are, if taken in isolation, consistent with Jiménez and Méndez's (1999) argument that implicit learning will be unaffected by distraction or interference, provided selective attention is paid to the predictive dimension. In other words, the system will not process input in a completely non-selective way, but instead associate those representations that are simultaneously activated in working memory (i.e., those that are the focus of attention). The implication is that as long as there is sufficient attentional capacity to perform the primary task, then learning will proceed normally. Any additional tasks that are to be performed while responding to the SRT task, no matter to what degree they require attentional resources, will not interfere with sequence learning. However, this claim is incompatible with the Shanks et al. study and the

experiment reported in Chapter 2, which found a detrimental effect on learning by introducing a resource-demanding secondary task. In the present context, the favoured conclusion is that the availability of working memory is critical for the instantiation of optimal sequence learning. To understand the importance of working memory for learning in the SRT task some recent work on attentional control is briefly described.

Lavie's load theory of attentional selection and control (Lavie, 2000, 2005; Lavie et al., 2004) proposes two mechanisms of attention involved in processing goal-relevant stimuli and ignoring irrelevant distractor stimuli. One is a passive perceptual selection mechanism that processes input until capacity is reached, and the other is an active mechanism of attentional/cognitive control that is involved in response selection, maintaining processing priorities, and determining the focus of attention. The distinction between perceptual and control mechanisms of attention is akin to the distinction made in the Introduction between input attention and central attention. The perceptual mechanism (input attention) will process irrelevant stimuli, provided sufficient capacity remains after goal-relevant stimuli have been processed. Therefore, increasing the perceptual load of the relevant task (by adding more stimuli or making the task harder) reduces capacity available for distractor processing, and thus *decreases* distractor processing. Conversely, increasing cognitive load (e.g., working memory load and/or dual-task co-ordination) results in *increased* distractor processing. This is because cognitive load reduces capacity available for cognitive control (which in turn decreases the efficiency of response selection), interferes with the maintenance of processing priorities, and, by disrupting the control of selective attention, allows poor defence against low-priority stimuli gaining control of behaviour (Lavie & de Fockert, 2005).

In the present experiments, the difficulty of selection was manipulated by introducing irrelevant distractors into the SRT display. Plainly, perceptual load was not high enough to prevent distractor perception, as evidenced by the increased RTs and negative priming (Lavie & Fox, 2000; Rees et al., 1997). Intuitively, one might expect that perception of such distractor stimuli in the context of the sequence learning task would impair learning by either depleting resources for target processing or by interfering with associative processes in working memory. However, no effect was observed. Application of Lavie's theory makes sense of this finding because it predicts that interference from perceived distractors can be minimised if the primary task induces low cognitive load (i.e., the absence of a secondary task or low working

memory load), which is precisely the condition created here. However, imposing a working memory load would predict that distractor interference increases because low-priority stimuli (e.g., symbols, tones, numbers held in working memory) may gain control of behaviour, as shown by the experiment in Chapter 2 and Shanks et al.'s (2005) study. It is thus argued here that attentional resources are required to ensure that goal-relevant processes in the SRT task - such as response selection, maintenance of processing priorities, attending to the contents of phenomenal awareness - are completed effectively and efficiently.

These results are now considered with respect to the implicit/explicit distinction. No measure was taken of 'implicit' learning as normally conceptualised (i.e., unconscious learning), because, following many previous arguments, it is posited that representations of what has been learned in implicit learning tasks are invariably conscious (Perruchet & Vinter, 2002; Shanks & St John, 1994). Indeed, there is evidence that learning in the probabilistic SRT task yields conscious sequence knowledge (Perruchet & Amorim, 1992; Shanks et al., 2003, 2005; Wilkinson & Shanks, 2004). It has been suggested that the content of phenomenal awareness during such experiments is representative of what has been attended, the focus of which has been the primary task and the on-line processing of input, particularly goal-relevant input (see Perruchet & Vinter, 2002). Additional cognitive load from a working memory task constitutes other goal-relevant, attended information in awareness, and thus, control of attention to the competing contents of consciousness may be compromised, resulting in less focus on learning relationships between contiguous stimuli from the SRT task. Both implicit and explicit learning are likely affected in this way. In contrast, implicit and explicit learning may be differentially impaired by disruption to the selection process. Indeed, there is evidence that memory recall in an explicit task is disrupted by divisions of attention at encoding, even when selective attention is paid to the target items (Isingrini, Vazou, & Leroy, 1995; Parkin & Russo, 1990), whereas the present experiments demonstrate that implicit learning is resistant to disruption of selection.

The basis of this dissociation may be best understood by considering the information processing properties of implicit and explicit learning (e.g., St. John & Shanks, 1997). In this view, implicit learning is deemed to operate slowly, incrementally, in a complex environment (see also Perruchet & Vinter, 2002), processing stimuli superficially without much regard for elaborative or intentional

strategies. Consequently, load on input attention is not likely to impair implicit learning, because the processes during encoding are extremely elementary, concerned primarily with the rapid processing of simple sensory features. The more complex operations of associating locations over several trials are performed in working memory and require central resources. Conversely, explicit learning often involves processing stimuli to a deeper level, which requires more time and elaboration, as well as hypothesis testing, where appropriate. Clearly, divisions of attention during encoding are likely to interfere with these processes.

Possibly one major criticism of the present study, and the conclusions drawn from it, is that sufficient evidence is not provided that load on input attention is innocuous only under conditions of low cognitive load. This conclusion would be upheld if distractors were shown to impair learning under conditions of working memory load. One approach therefore is to study the effects of distractors when the primary task is performed under conditions of high load. Unfortunately, learning scores under conditions of working memory load are often very small (approximately 10 milliseconds, Shanks et al., 2005; but see Jiménez and Méndez, 1999), and thus leave little scope for detecting a detrimental impact from added distractors. Chapters 5 and 6 consider in more detail possible methods for investigating the effects on SRT learning of perceptual load and working memory load simultaneously.

Three other potential causes of the absence of an effect of selection difficulty in the present experiments are acknowledged. First, it is possible that training was not long enough. Jiménez & Vázquez (2005) have recently shown that during probabilistic sequence learning attention-demanding tasks only become detrimental to learning during later stages of training (when, they argue, explicit knowledge of the sequence is acquired). Perhaps decrements in learning would have been found had training been considerably longer. This possibility is not, however, consistent with studies that show dual-task effects after short amounts of training (Shanks et al., 2005), nor with some work presented in Chapters 4 and 5 of this thesis. A second possibility is that the selection task in the present experiments was not sufficiently demanding. Perhaps if participants were required to make a response to the distractor, yet still select the target, learning would be attenuated. In support of this is the finding that implicit *memory* is impaired by having to respond to two dimensions of a word (i.e., colour and phonology; see Mulligan & Hornstein, 2000). One problem with such a procedure is that it resembles dual tasking, and this has been shown to have effects similar to loading

working memory (Lavie et al., 2004) and task switching (Pashler, 1998). Alternatively, the filtering cost could be made more demanding by making the selection process far more difficult. Although this was done in Experiment 5 (by selection being based on category discrimination), it remains to be seen whether, for example, selecting the target amongst several very similar stimuli, or amongst members of complex categories, would yield different results. Indeed, if selectional demands drew considerably on working memory resources – say, if participants had to select the highest number from a pair – then impairment in sequence learning should be anticipated. This chapter has reported that increases in selection difficulty do not impair implicit sequence learning, yet one must not assume that different selection tasks, or different implicit learning tasks, will invariably yield the same conclusion.

Finally, it was mentioned earlier that some authors have argued that sequence learning depends heavily on motor learning, and it follows from this that using perceptual distractors is an inappropriate manipulation for studying the effects of input attention on implicit learning. Although the premise of this view was disputed in the introduction to this chapter by presenting a wealth of evidence in support of a major contribution of perceptual learning in sequence learning, it would certainly be interesting to extend the present results by asking whether similar findings are obtained when participants learn by observation. Under such observational conditions, motor learning is eliminated.

# Chapter 4

## Incidental Learning of Unattended Stimuli

This chapter presents some pilot data and provides the foundation for Chapter 5. The experiments included here were motivated by the question of whether it is possible to learn about stimulus regularities that are not the focus of selective attention. In short, two experiments using probabilistic sequences examined whether participants could learn a target sequence in the SRT task and simultaneously learn something about an incidental secondary sequence.

Currently, there is some discrepancy on this issue. One view, held by Jiménez and Méndez (1999), is unequivocal in its assertion that selective attention is necessary for learning, including unconscious or implicit learning. According to Jiménez and Méndez, if additional regularities in a sequence learning task are presented simultaneously with the target sequence, then participants will only learn about these secondary regularities if they pay selective attention to them. Their experiment, which was reviewed in detail in both Chapters 1 and 2, showed that participants in a single task group did not learn about another relevant (but not selectively attended) predictive dimension even though it was beneficial to learning about sequence locations in the SRT task. The dual task group on the other hand, who had to attend to the other relevant dimension in order to perform the concurrent task, did learn about this secondary sequence. Jiménez and Méndez concluded that secondary stimulus regularities are only learned if they are selectively attended, even though they may be relevant to the primary task.

In contrast, Cock et al. (2002) provided evidence that a secondary unattended sequence can be learned. In their experiments, participants were presented with two differently coloured stimuli each following a different sequence. The aim of the task was to respond to one of the stimuli (the target) and ignore the other (the distractor). The results appeared to suggest that participants learned both about the target and the distractor sequences.

These conflicting results might best be understood by considering some recent evidence using the contextual cueing paradigm. In this task, which was reviewed in Chapter 1, participants are required to search for a target amongst distractor displays (visual search), but crucially, they appear to "implicitly" develop sensitivity to the location of targets (measured as decreasing RTs to targets across practice blocks) that

are presented at consistent locations in repetitive target-distractor displays. Interestingly, a new study using this task (Jiang & Leung, 2005) has shown that participants can also acquire sensitivity to an ignored distractor set, thus suggesting that implicit learning of ignored visual context can occur in the absence of selective attention to the elements of the ignored display. The most important finding of that study however was that learning of the ignored distractor set was only expressed when it subsequently became the focus of attention, despite this learning having been acquired outside the focus of attention. It is thus possible that learning of the secondary sequence in the Jiménez and Méndez (1999) study did occur for the single-task group, but that it was not expressed because the secondary sequence never became the focus of attention. The Cock et al. (2002) study on the other hand did switch attention to the previously ignored sequence, and subsequently, learning of this sequence was found to have occurred.

However, despite the apparent positive result obtained by Cock et al., the details of their methodology make it difficult to draw any firm conclusions. In their Experiments 2 and 3 the target and distractor sequences were correlated such that on each trial there was a relationship between the location of the distractor stimulus and the location of the target stimulus. Because of this, it is unclear whether participants were learning two sequences simultaneously, or just one compound sequence composed of two stimuli (Schmidtke & Heuer, 1997). However, a greater concern is that, by presenting 2 sequences that are correlated, it becomes not only easier to learn about the secondary sequence, but also it becomes more relevant. Moreover, the sequences were very simple (merely 6 deterministic transitions each), and the stimuli used for the separate sequences were clearly distinguishable (being blue and red). Bearing in mind these preceding points, is it not entirely likely that participants were paying selective attention to the to-be-ignored events? If this were the case, then the question of whether selective attention is necessary for implicit sequence learning has not been fully addressed by the Cock et al. (2002) study.

As the above discussion makes clear, the issue about whether selective attention is necessary for implicit learning is far from resolved. The experiments in this chapter provide some new data on this topic, and highlight some important methodological difficulties in this area of research.

## 4.1 Experiment 6

This experiment was designed to find out whether participants learn about a distractor sequence that is presented simultaneously with a target sequence. It has a number of features that are similar to the Cock et al. set-up. First, to aid identification, the target and distractor stimuli were different colours. Second, the target sequence and the distractor sequence were presented in the same array of screen locations. Third, participants responded to the location of the target whilst attempting to ignore the distractor sequence. Therefore, unlike the Jiménez and Méndez experiments described above, there is no requirement to attend to the secondary sequence, yet participants were clearly able to see that another dimension was present. In Jiménez and Méndez's study it was not at all clear that there was another dimension present because the secondary sequence was presented through the changes in target identity between trials. Participants in the single task group could simply ignore this and just respond to the target location on each trial. Thus, the experiments presented in this chapter are chiefly concerned with investigating incidental learning of salient, yet to-be-ignored, distractor sequences.

Unlike the sequences in the Cock et al. (2002) study, in the present experiment there was no correlation between the target and distractor sequences. Therefore, to learn anything about the distractor sequence whilst learning the target sequence required participants to learn about two independent sequences simultaneously. Additionally, the sequences in the present experiment were longer (12 items) and more complex than those used by Cock et al., hence, they were more difficult to learn. With more attentional capacity being consumed by learning the noisy sequences (Howard & Howard, 2001), it is thus less likely that participants would be able to devote selective attention to the secondary stimuli (Lavie & de Fockert, 2005).

To ensure that latent learning (should it occur) of the distractor sequence was expressed, testing of the ignored sequence took place under conditions that required the ignored sequence to be the focus of attention (Jiang & Leung, 2005). This was achieved by removing the distractor for the final 2 blocks of the SRT task, and then the target followed one of three sequences: (1) the target sequence during the training blocks of the SRT task; (2) the distractor sequence during the training blocks of the SRT task; or, (3) a completely new sequence. Each of three groups was assigned to one of the conditions. The purpose of including a group who react to a new sequence was to provide a measure of how much is learned about a sequence in 2 blocks (the testing

blocks). If the group who react to the distractor sequence during the testing blocks show a greater degree of learning than the new group, then it is likely that they learned about the distractor sequence during the training blocks when they were responding to the target sequence. The group who continued responding to the target sequence during the test blocks provided a measure of learning for the attended sequence over all blocks of the SRT task. This is essential in order to determine how much is learned about the ignored sequence in comparison with the attended sequence. During the test blocks the target identity changed to that of a black cross. The purpose of this was to achieve parity of testing conditions between the three groups.

The above method of assessing learning of the different sequences requires that participants respond directly to their assigned sequence without distraction from competing stimuli. In the present experiments it is important to remove any competition during the testing stage between target and distractor stimuli because of the method adopted here for measuring learning. Unlike Cock et al., who used deterministic sequences, the present experiments use probabilistic sequences. Learning of these sequences is measured by computing the RT difference between probable and improbable trials within blocks, and not, as is the case with the deterministic sequences used by Cock et al., by calculating the differences between absolute RTs between sequences across blocks. For this reason, it is not clear whether the negative priming analyses employed by Cock et al. are relevant here. In their experiments, learning of the ignored sequence was detected by measuring whether RTs to targets that followed a previously ignored sequence were slower during the test phase than were RTs to the control sequences (i.e., responses to these targets were negatively primed). But the probabilistic sequences used here are less likely to facilitate the development of negative priming at the sequence level. For instance, negative priming beyond single locations is not likely to occur when the sequences are uncorrelated because the lack of regularity between them would tend to preclude the build-up of a patterned set of inhibitory responses. Instead, a more likely outcome if learning does occur is that participants will become sensitive to the underlying statistical nature of the distractor sequence, and that this learning will be facilitated, or 'released' when this sequence becomes the focus of attention in the test phase (Jiang & Leung, 2005). Nevertheless, the experimental design does allow negative priming to be measured, both for single locations in the training phase, and for the previously unattended sequence in the test phase. According to some theories of negative priming, repeated exposure to a

previously ignored stimulus leads to positive rather than negative priming, but only in the absence of competition from distractors (i.e., when there is no load on selection mechanisms; Milliken, Joordens, Merikle, & Seiffert, 1998). For this reason, as mentioned above, just a single stimulus was presented in the test phase of this experiment.

In summary, the present experiment was designed to investigate learning of an unattended distractor sequence in a probabilistic sequence learning task. Testing conditions were identical for all three groups, and learning was measured by responding to the location of the target sequence during testing blocks presented after the training session. During the testing blocks, the target sequence was different for each group, being either the previously attended sequence, the unattended sequence, or a new sequence.

### 4.1.1 Method

#### *4.1.1.1 Participants*

Ninety-nine individuals (79 women and 20 men, mean age = 20.2 years, range = 17-40) completed the experiment as part of a University College London undergraduate psychology class. They were randomly assigned to one of three groups, attended ($n$=29), unattended ($n$=37) or new ($n$=33). They were not remunerated.

#### *4.1.1.2 Apparatus and Materials*

These were identical to those used in Chapter 3, except for the following: Six 12-item second-order conditional (SOC) sequences were used in this experiment; these are referred to as SOC1 (2-1-2-3-4-1-3-2-4-3-1-4), SOC2 (4-1-4-3-2-1-3-4-2-3-1-2), SOC3 (3-1-3-2-4-1-2-3-4-2-1-4), SOC4 (4-1-4-2-3-1-2-4-3-2-1-3), SOC5 (1-2-1-3-4-2-3-1-4-3-2-4), and SOC6 (4-2-4-3-1-2-3-4-1-3-2-1). The training sequences and the alternative sequences were paired as follows: SOC1/SOC2, SOC3/SOC4, SOC5/SOC6.

Three stimuli were presented in the SRT task: a red circle 3 mm in diameter, a green square with sides of length 5 mm, and a black cross that was approximately the same size as the square.

#### *4.1.1.3 Design and Procedure*

There were three groups in this experiment: attended, unattended, and new. During the SRT task all groups performed 10 blocks of 100 trials each. For blocks 1-8

all participants responded to the location of the target stimulus (red circle) while simultaneously having to ignore the unattended stimulus (green square). On blocks 9 and 10 the unattended stimulus was removed and the identity of the target stimulus was altered, becoming a black cross. The order of screen locations at which the target and distractor appeared was consistent with probabilistic SOC sequences. On each trial, and for all sequences, a probable trial occurred with probability .85 and an improbable trial with probability .15. The target sequence and the distractor sequence were different and uncorrelated. For the attended group the target sequence was identical for all 10 blocks of the SRT task. For the unattended and new groups, the target sequence on blocks 9 and 10 was different from the target sequence on blocks 1-8. For the unattended group the target sequence on blocks 9 and 10 was identical to that of the unattended sequence on blocks 1-8. For the new group the target sequence on blocks 9 and 10 was a novel sequence. The three pairs of SOC sequences (see *Materials*) were randomly assigned to conditions. For each sequence in a sequence pair the assignment as either the training or the alternative sequence was random. The design is shown in Table 4.1.

Table 4.1. The experimental design of Experiment 6. SOC refers to the second-order conditional sequences used.

| Experimental Group | Sequence on Blocks 1-8 | | Sequence on Blocks 9 & 10 (black cross only) |
|---|---|---|---|
| | Target (red circle) | Distractor (green square) | |
| Attended | SOC A/B | SOC C/D | SOC A/B |
| Unattended | SOC C/D | SOC A/B | SOC A/B |
| New | SOC C/D | SOC E/F | SOC A/B |

The purpose of the experiment was to investigate whether participants learn the unattended sequence. If they do, then learning scores for the unattended group on blocks 9 and 10 will be greater than those for the new group. If, on the other hand, learning is restricted to the attended sequence, then learning scores for the unattended and new groups on blocks 9 and 10 will be the same.

### 4.1.2 Results

The alpha level is set at .05, two-tailed, throughout, unless otherwise stated. RTs on error trials were excluded from the analyses and RT data were collapsed across all

training sequences. An error rate criterion was adopted such that participants with a mean error rate of greater than 10% (across all blocks and trials) were excluded from the analyses ($n$ = 17). The error rates were exceptionally high in some cases, possibly because this experiment was run as an undergraduate laboratory class, and thus some participants took less care than they might have under more controlled conditions. The total number of participants included in the analyses was 82 (attended $n$ = 24, unattended $n$ = 31, new $n$ = 27).

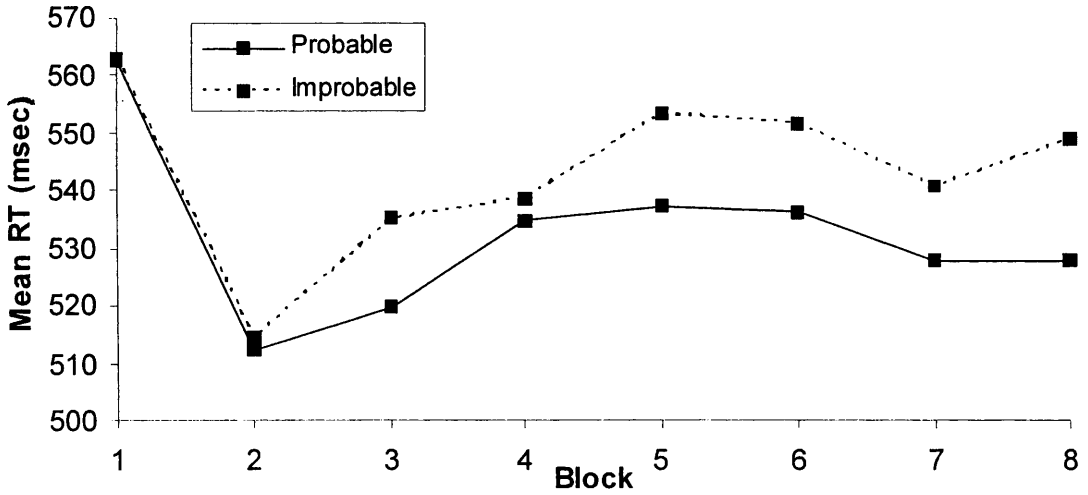


Figure 4.1 Mean RTs for probable and improbable targets across the training blocks (1-8) of Experiment 6.

*4.1.2.1 Response-Time Data for Training Blocks (1-8)*

Figure 4.1 shows the mean RT data across the training blocks (1-8) of the SRT task with the three groups collapsed (due to functionally identical training conditions) and probable and improbable target data plotted separately. Participants' mean RT data were entered into a 3-way mixed-model analysis of variance (ANOVA) with the within-participants factors of Block (8 levels) and Target Probability (2 levels, probable and improbable), and the between-participants factor of Group (3 levels, attended, unattended, and new). There was a highly significant main effect of Target, $F(1, 79)$ = 18.97, $MSE$ = 2, 024, $p < .01$, yet no Target × Group interaction, $F<1$. These results suggest that all three groups learned the training sequence equally. There was also a Block × Target interaction, $F(7, 553)$ = 2.06, $MSE$ = 1, 114, which reflects the fact that learning emerged with practice. There was no main effect of group, and no 3-way

interaction, *F*s < 1, thus all groups performed and learned the attended sequence equally.

*4.1.2.2 Error Data for Training Blocks (1-8)*

Error rates for target localisation errors on blocks 1-8 were entered into an ANOVA with Block (1-8) and Target Probability (probable and improbable) as within participants factors, and Group (attended, unattended, and new) as a between participants factor. There was a significant main effect of Target Probability, $F(1, 79) = 10.43$, $MSE = 28$, $p < .01$, which indicates that participants made proportionally more errors on improbable trials ($M = 14.0\%$) than on probable trials ($M = 8.3\%$). Such a result is reflective of the anticipatory responses towards sequence consistent locations that are acquired during sequence learning. No group effect or interactions emerged (*F*s < 1.19), suggesting that error rates were equal across the groups.

*4.1.2.3 Negative Priming*

Negative priming was measured in this experiment because, as it played an important role in the Cock et al. (2002) study, it is important to determine the extent of distractor processing. Target trials where the current target location was occupied by the distractor on the immediately preceding trial may be slower due to negative priming. A comparison between negatively primed (NP) trials ($M = 567$ milliseconds, $SE = 10.45$) and non-NP trials ($M = 551$ milliseconds, $SE = 10.22$) reveals this to be the case, $t(81) = 7.32$, $p < .01$.

*4.1.2.4 Learning Scores*

The principal data of interest were the differences between RTs for probable and improbable targets on the test blocks (i.e., Blocks 9 and 10). The difference scores were calculated by subtracting the block mean probable RT from the block mean improbable RT for each participant and then averaging over all participants for that group. The difference scores for each group on blocks 9 and 10 are shown in Figure 4.2.

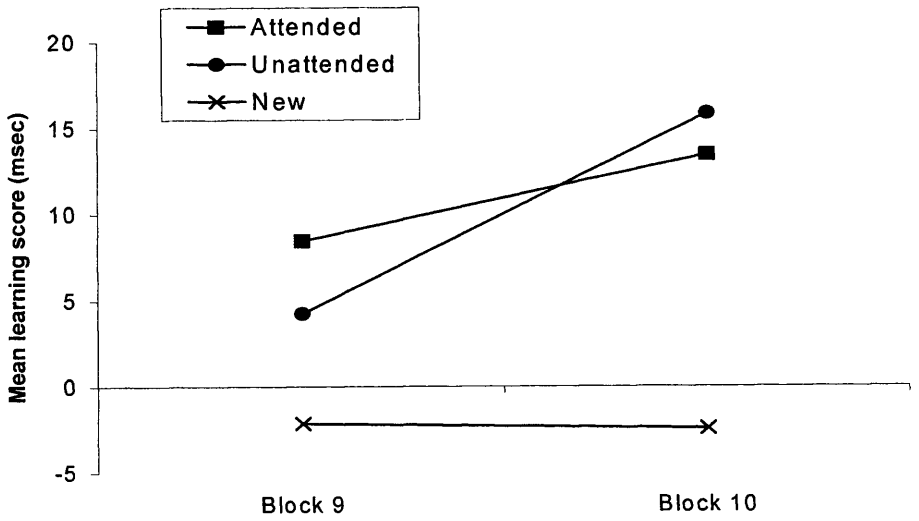


Figure 4.2 Mean learning scores for the *attended, unattended,* and *new* groups on the testing blocks (9 and 10) of Experiment 6.

Unexpectedly, there is a large decrease in learning scores (from 20 milliseconds to 8 milliseconds) for the attended group between the last training block (Block 8) and the first testing block (Block 9). Recall that the attended group continued responding to the same sequence during this switch, yet the distractor was removed and the target stimulus changed identity (from a red circle to a black cross). It is likely that this decrement in the expression of learning was due to the change in stimulus identity and not to the removal of the distractor (Chapter 3 supports this view). Unfortunately, this short-term abolishment of learning (it increased somewhat on Block 10) makes the results difficult to interpret. Therefore, with this in mind, the focus is on the testing data from Block 10.

A one-way ANOVA on the block 10 learning scores for the three groups (attended $M$ = 13 milliseconds, $SE$ = 8.02; unattended $M$ = 16 milliseconds, $SE$ = 5.11; new $M$ = -2 milliseconds, $SE$ = 8.03) suggested that all groups learned the sequence equally, $F(2, 79) = 2.05$, $MSE$ = 1, 331, $p = .14$. However, a comparison between the unattended and new groups revealed a statistically significant difference in learning scores, $t(56) = 1.97$, $p < .05$, one-tailed. Moreover, one-sample t-tests showed that both the attended group, $t(23) = 1.68$, $p = .05$, and the unattended group, $t(30) = 3.10$, $p < .01$, one-tailed, obtained learning scores greater than zero, but that the new group did not, $t(26) = 1.31$, *ns*. In contrast, there was no difference in learning scores between the attended and new groups, $t(49) = 1.40$, *ns*.

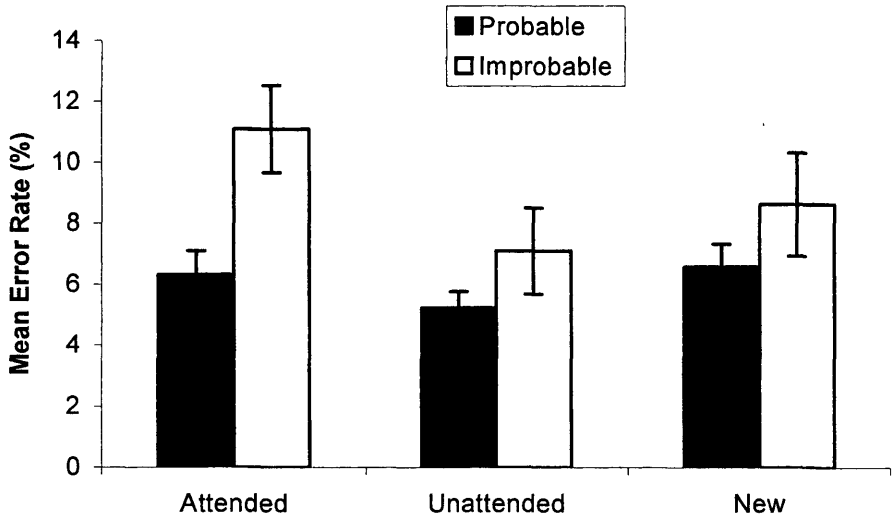


Figure 4.3 Mean percentage error rates for probable and improbable targets for the *attended, unattended,* and *new* groups on test Block 10 of Experiment 6.

*4.1.2.5 Error Data for Test Block 10*

Figure 4.3 shows the mean error rates for Block 10. Mean error rates from block 10 were entered into a two-way ANOVA with the within participants factor of Target (probable and improbable) and the between participants factor of Group (attended, unattended, and new). There was a main effect of Target, $F(1, 79) = 9.42$, $MSE = 36$, $p < .01$, yet no main effect of Group and no Group × Target interaction, $Fs < 2.26$, suggesting that target localisation errors were consistent across groups and thus that any learning differences were not manifest in this way. However, on comparisons between errors for probable and improbable trials, the attended group show a statistically significant difference, $t(23) = 3.03$, $p < .01$, with improbable errors ($M = 11.11\%$, $SE = 1.43$) being higher than probable errors ($M = 6.32\%$, $SE = 0.78$), whereas the unattended and new groups did not show this difference between probable and improbable errors on block 10, $t(30) = 1.21$, *ns*, and $t(26) = 1.18$, *ns*, respectively.

4.1.3 Discussion

The pattern of results reported for block 10 appears to suggest that both the attended and the unattended group showed some weak evidence for sequence learning, and that the new group did not. Similar to the findings of Cock et al. (2002), there was evidence that negative priming for single locations operated during the training phase of

the experiment. However, in contrast to the findings by Cock et al., there was no evidence from the RT data collected during the test phase that negative priming had occurred at the sequence level, which suggests that participants had not learned to inhibit responses to unattended sequence transitions. Instead, the learning scores revealed that probable locations in the previously unattended sequence were *positively* primed when this sequence became the target sequence in the test phase. This result supports the claim made by Jiang and Leung (2005) that implicit learning of ignored visual context can be expressed (by facilitating responses) when the previously unattended information is subsequently attended. Unfortunately, the results from this experiment are somewhat unclear because of the apparent short-term expression decrement that arose on Block 9 due to the change in stimulus identity.

## 4.2 Experiment 7

The previous experiment investigated whether participants learn about an unattended distractor sequence. There was some weak evidence that they may, but the results were far from conclusive. One major problem seems to be that the change in stimulus identity between the training and test blocks resulted in an inability to immediately apply sequence knowledge to a different stimulus. Therefore, the present experiment uses the same stimulus throughout the training and test blocks and for the target and distractor stimuli. This is achieved by having two rows of locations, one that presents the target sequence and one that presents the distractor sequence. The stimulus is identical in each row. During training, participants responded to the target sequence in the bottom row whilst ignoring the distractor sequence in the top row. During testing, participants responded to the target sequence in the top row and ignored the distractor in the bottom row. As in Experiment 6, the target sequence in the testing blocks followed either the target sequence from the training blocks, or the distractor sequence from the training blocks, or a new sequence. The logic was identical to that of Experiment 6.

Another motivation for presenting the target and distractor sequences at separate arrays of locations was to avoid the occurrence of negative priming. Negative priming, although not evident at the sequence level in the previous experiment, did occur for single locations. It is possible that this interfered with either the learning or the expression of the distractor sequence (and possibly the target sequence). The present experiment therefore sought to avoid this potential confound. Another reason for attempting to reduce the occurrence of negative priming is that according to influential

accounts of negative priming (e.g., Houghton & Tipper, 1994), this phenomenon actually reflects the operation of an inhibitory mechanism of selective attention. Thus, in this view, evidence for the existence of negative priming to ignored events supports the notion that these events were in some sense attended. Therefore, by minimising these effects, the following experiment was more suitable for investigating the role of selective attention in implicit sequence learning. Specifically, if learning about a distractor sequence occurs in the absence of any negative priming effects of single locations, then the evidence would be more convincingly supportive of the view that implicit learning can proceed independently of the need to selectively attend to sequence elements.

### 4.2.1 Method

#### *4.2.1.1 Participants*

Ninety-seven individuals (83 women and 14 men, mean age = 19.7 years, range = 18-38) completed the experiment as part of a University College London undergraduate psychology laboratory class. They were randomly assigned to one of three groups, attended ($n$=29), unattended ($n$=36), or new ($n$=32). They were not remunerated.

#### *4.2.1.2 Apparatus*

Four boxes (8 $mm^2$, with white backgrounds) were set centrally against a grey background, and aligned horizontally on the screen with a spacing of 32 mm between each. The positions of the boxes will be referred to as locations 1-4 from left to right, respectively. 2.5 millimetres above this row of boxes was another row of boxes, identical to the one below (see screen shot, Figure 5.1, in Chapter 5).

#### *4.2.1.3 Materials*

The sequences and their pairings were identical to those of Experiment 1. Two red circles (3mm in diameter) were used as stimuli.

#### *4.2.1.4 Design and Procedure*

Participants performed the SRT task for 10 blocks. The design is shown in Table 4.2.

Table 4.2. The experimental design of Experiment 7. *SOC* refers to the second-order conditional sequences used.

| Experimental Group | Sequence on Blocks 1-8 | | Sequence on Blocks 9 & 10 | |
|---|---|---|---|---|
| | Bottom Row (target) | Top Row (distractor) | Bottom Row (distractor) | Top Row (target) |
| Attended | SOC A/B | SOC C/D | SOC C/D | SOC A/B |
| Unattended | SOC A/B | SOC C/D | SOC A/B | SOC C/D |
| New | SOC A/B | SOC C/D | SOC A/B | SOC E/F |

For the first 8 blocks participants responded to the location of the target that appeared in the bottom row of boxes. A distractor stimulus, which participants were required to ignore, appeared in the top row of boxes. Both stimuli followed different probabilistic SOC sequences. On Blocks 9 and 10 participants were required to switch from responding to the target appearing in the bottom row of boxes and respond to the location of the stimulus in the top row. The stimulus in the bottom row continued to appear at locations consistent with the sequence it had followed during Blocks 1-8. The crucial manipulation was that the sequence on the top row during Blocks 9 and 10 could either follow the same sequence as it had followed during Blocks 1-8 (the *Unattended* group), or it could follow the same sequence as the target had followed on the bottom row during Blocks 1-8 (the *Attended* group), or it could follow a sequence that had not been seen before (the *New* group). The assignment of sequences to conditions and target/alternative sequence pairs was done in the same way as in Experiment 6.

This design allowed an assessment of whether participants learned anything about a sequence which was not the focus of attention (i.e., it was not selectively attended to), and to what extent learning of the unattended sequence compared with the degree of learning of the attended sequence. If participants learn something about the unattended sequence during Blocks 1-8, then, on Blocks 9 and 10 the learning score for the unattended group should be greater than that of the new group. Furthermore, if the attended group learn more about the sequence than the unattended group then learning scores for the attended group participants should be greater than those for the unattended and new groups. The new group provided a baseline for what is learned about a sequence over 2 blocks having already performed 8 blocks of the SRT task.

4.2.2 Results

The total number of participants included in the analyses after excluding those ($n$ = 25) who had high error rates (same criterion as Experiment 6) was 72 (attended $n$ = 23, unattended $n$ = 23, new $n$ = 26).

*4.2.2.1 Response-Time Data for Training Blocks (1-8)*

Figure 4.4 shows the mean RT data for Blocks 1 – 8. The data were entered into an ANOVA with Block, Target and Group as factors. Most importantly there was a highly significant main effect of Target Probability, $F(1, 69) = 35.46$, $MSE = 2{,}174$, $p < .01$, and a Block × Target interaction, $F(7, 483) = 2.25$, $MSE = 1{,}452$, showing that sequence learning improved with practice. There were no effects of group or interactions with the factor of group ($F$s < 1.39), thus, as with Experiment 6, there were no differences in performance or learning of the attended sequence between the three groups.

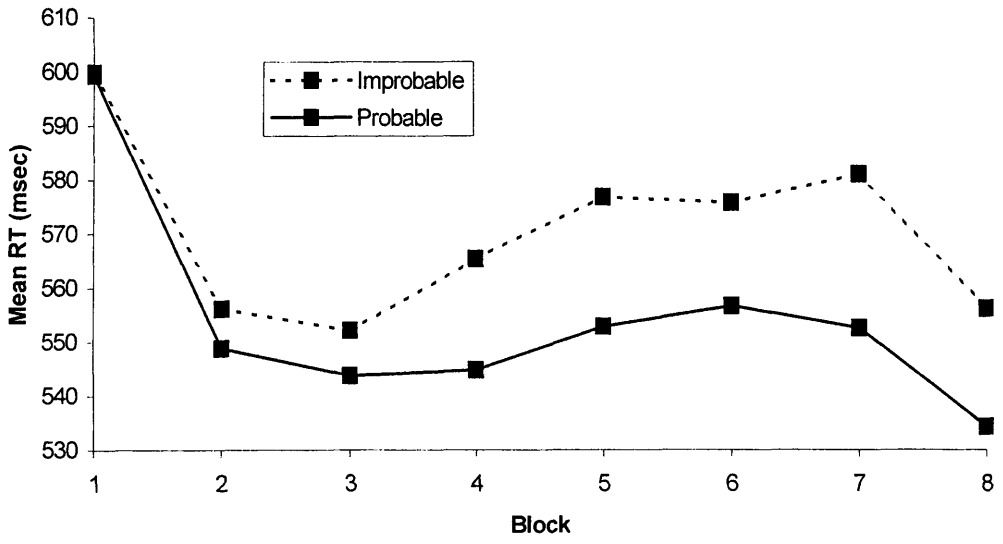


Figure 4.4 Mean RTs for probable and improbable targets across the training blocks (1-8) of Experiment 7.

*4.2.2.2 Error Data for Training Blocks (1-8)*

Mean target localisation error rates were entered into an ANOVA with Block, Target and Group as factors. As with the RT data, there was a highly significant main effect of target (proportionally more errors to the improbable targets, $M$ = 13.0 %, than

the probable targets, $M$ = 8.0 %), $F(1, 69) = 12.40$, $MSE = 36$, $p < .01$, and a Target × Block interaction, $F(7, 483) = 3.14$, $MSE = 24$, $p < .01$. There were also no group effects ($F$s < 1). Consistent with the RT data, these results demonstrate that learning of the attended sequence improved with practise and was equal for all groups.

*4.2.2.3 Negative Priming*

Negative priming (i.e., in this case trials where the sequence location, e.g., 1-4, of the target in the bottom row was the same as the location of the distractor in the top row on the immediately preceding trial) was unlikely to be as important in this experiment compared with Experiment 7 because distractor locations were spatially separate from the target locations. Indeed, there was 1 millisecond difference between NP trials (581 milliseconds) and non-NP trials (582 milliseconds). No further consideration was given to this issue.

*4.2.2.4 Learning Scores*

Figure 4.5 shows the mean difference scores on Blocks 9 and 10 for all groups. It is clear from the graph that the attended group has numerically higher learning scores than do the new or unattended groups. The learning scores appear to increase across the 2 testing blocks. The data were entered into an ANOVA with the factors of Block and Group. The effect of block approached significance, $F(1, 69) = 3.88$, $MSE = 1{,}467$, $p = .053$, whereas the Block × Group interaction did not ($F < 1$). The group factor approached significance, $F(1, 69) = 2.91$, $MSE = 2{,}339$, $p = .06$. These results suggest that there may, as there was in Experiment 6, be a decrement in the expression of learning on Block 9 which is due to the switch in response location from the bottom to the top row. An ANOVA on the mean learning scores collapsed across blocks approached significance, $F(2, 71) = 2.91$, $MSE = 1{,}170$, $p = .06$ . Further comparisons revealed that the learning score of the attended group differed from that of the unattended group, $t(44) = 2.15$, and the new group, $t(47) = 1.90$, one-tailed. The learning scores for the unattended and new groups did not differ $t((47) = 1.16$, *ns*. Moreover, the attended group obtained a learning score greater than zero, $t(22) = 3.34$, $p < .01$, whereas the unattended and new groups did not, $t(22) = 0.53$, *ns*, $t(25) = 1.21$, *ns*.

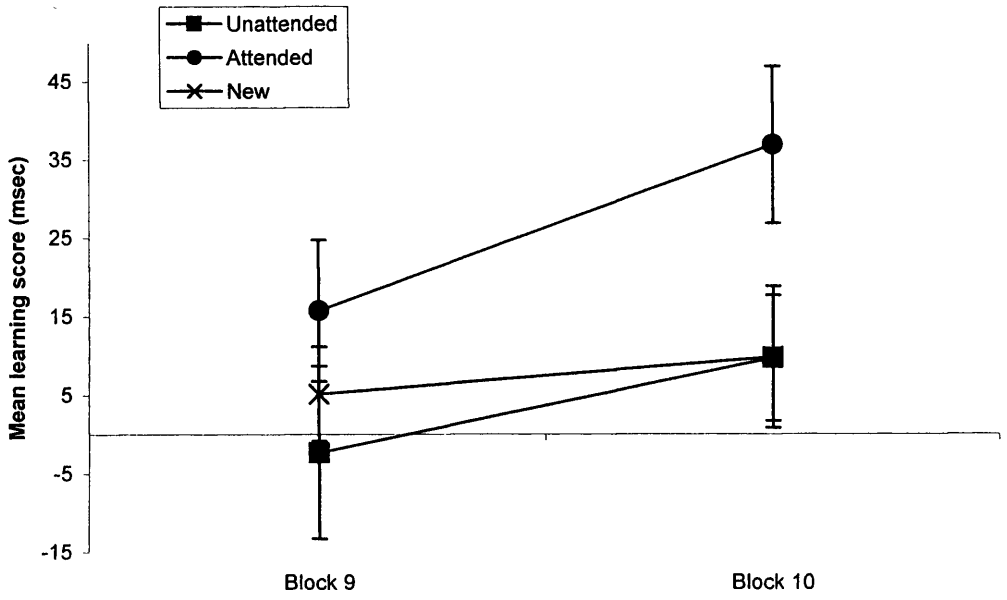


Figure 4.5 Mean learning scores for the *attended, unattended,* and *new* groups on the testing blocks (9 and 10) of Experiment 7.

### *4.2.2.5 Error Data for Testing Blocks (9 & 10)*

Figure 4.6 shows the mean error rates for improbable and probable trials collapsed across Blocks 9 and 10.

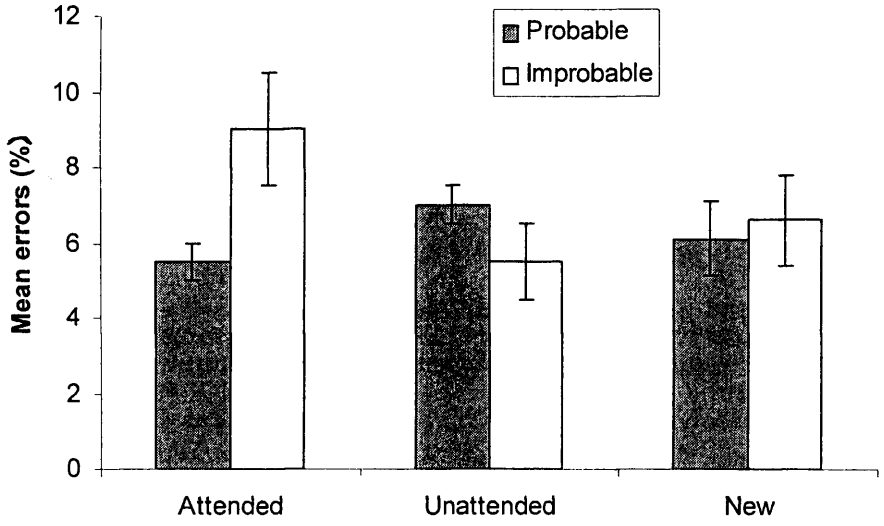


Figure 4.6 Mean errors for probable and improbable targets for the *unattended, attended,* and *new* groups collapsed across the test Blocks 9 and 10 of Experiment 7.

An ANOVA on the data, with Block, Target and Group as factors, reveals a Target × Group interaction, $F(2, 69) = 6.27$, $MSE = 27$, $p < .01$. No other effects or interactions reached significance ($F$s < 2.77). It is clear from the graph that for the attended group improbable errors are higher than probable errors, whereas this difference is not so pronounced for the other groups. Such a finding is consistent with the RT data that the attended group shows evidence of sequence learning on Blocks 9 and 10 whereas the unattended and new groups do not.

### 4.2.3 Discussion

Overall, these results show that participants do not learn about an unattended sequence. The experiment was an improvement on Experiment 6, because there was clearer evidence of learning of the attended sequence.

## 4.3 Discussion of Experiments 6 and 7

The two experiments reported in this chapter were designed to investigate learning of ignored distractor sequences. In the experiments, participants were instructed to respond to a target sequence during the training blocks of a probabilistic SRT task. They were simultaneously presented with a distractor sequence, which they were to ignore. The target and distractor sequences were not correlated. Thus, learning the distractor sequence would demonstrate incidental learning of an unattended sequence.

Experiment 6 provides some weak evidence that distractor sequences can be learned. However, it is unclear to what degree the distractor sequence was learned in comparison with the attended sequence because the expression of learning of the attended sequence was impaired somewhat by a change in stimulus identity between training and test. This problem was largely overcome in Experiment 7 (by using the same stimuli throughout). However, there was no evidence of unattended sequence learning in the latter experiment. There are several possibilities for these discrepant results. First, in Experiment 6, the distractor stimuli appeared at the same array of locations as did the target stimuli. This would make the distractor sequence more difficult to ignore (at a purely perceptual level). Second, in order to respond correctly to the target sequence, participants would have needed to inhibit responses to the competing distractor stimuli. Indeed, the negative priming data showed that this inhibition was prevalent, and thus participants were not entirely ignoring the distractor

stimuli. Finally, the identity of the targets and distractors were different from each other, and each was distinctive. Collectively, the above considerations point strongly to the notion that the distractor sequence in Experiment 6 was not an "unattended" sequence, and in fact, seems likely to have been allocated a considerable amount of attention. These same criticisms apply equally to Cock et al.'s (2002) study. Conversely, the distractor sequence in Experiment 7 is far less likely to have been successfully vying for attention. Recall that it was presented at a separate set of locations, and that the distractor stimulus was identical to that of the target stimulus. Under these conditions it is much easier to ignore the distractor sequence and focus attention exclusively on the target. The lack of negative priming in Experiment 7 supports this view that the distractor sequence was unattended.

Considering the above, it appears that truly unattended sequences are not learned. The present results therefore support Jiménez and Méndez's (1999) view that selective attention is necessary for implicit sequence learning. In contrast, the possibility that an ignored sequence can be learned, as the Cock et al. (2002) study seemed to show, has not been supported.

# Chapter 5

## Attention Modulates the Learning of Multiple Contingencies

Chapter 4 presented some pilot data on the role of selective attention in implicit sequence learning. The experiments included in that chapter examined whether participants could learn about a distractor sequence that was not the focus of attention, but the results were somewhat inconclusive. The present chapter therefore broadens the investigation to look more closely at the possibility of learning two sequences simultaneously, and takes a slightly different approach to understanding the role of attention in this process.

The study by Jiang and Leung (2005) described in Chapter 4 reported that participants can learn about a secondary unattended set of distractors in an implicit learning task (contextual cueing), and that this latent learning can be "released" – having a facilitatory effect on performance – when the unattended set subsequently becomes the focus of attention. Their result suggests that the attentional system does not filter out irrelevant information early in the perceptual stream, but instead processes all the available input to a high level. In this view, the attentional bottleneck operates at the response stage of processing, serving to guide behaviour only once semantic information has been extracted from sensory input. However, this late selection view of attention (e.g., Deutsch & Deutsch, 1963) has recently been challenged by Lachter et al. (2004), who convincingly argue that Broadbent's (1958) original conception of attention as an early selective filter is best able to accommodate the existing evidence. By considering a vast body of previous work and by introducing five new experiments, Lachter et al. concluded that in the absence of attention, only the simplest sensory features (e.g., contrast, motion, colour) of unattended stimuli are represented in early visual processing, and that identification (i.e., semantic processing or object representations) of stimuli only occurs once they have been selectively attended. In the instances where experiments have reported processing of unattended items, Lachter et al. argue, there is evidence of attentional spillage: spare processing capacity that is available after processing attended items is able to be devoted to processing of unattended items. If this account is correct, then it raises the possibility that the result obtained by Jiang and Leung (2005) came about because spare attentional capacity not consumed by processing of the attended distractor set, spilled over to allow processing

of the unattended set. Similarly, the notion of attentional spillage explains why Cock et al. (2002) observed learning of a secondary sequence; and undoubtedly this is why strong negative priming effects, themselves evidence of attentional operations (Houghton & Tipper, 1994), were prevalent in that study. From this perspective, it cannot be claimed that implicit learning of supposedly ignored distractors may occur in the absence of attention.

Lavie's (e.g., 1995, 2005) perceptual load theory of attention, which was described in Chapters 1 and 3, has been influential in helping resolve the early versus late selection debate. Briefly, Lavie argues that under conditions of low perceptual load (where only a few perceptual items need to be processed for a relevant task), excess capacity not consumed by the relevant task is available to spill over to processing irrelevant distractor items (i.e., late selection). Conversely, under conditions of high perceptual load (many items in the relevant task), all attentional capacity is used up by the relevant task, and thus there is none available to spill over to processing irrelevant distractors (early selection). Thus, whether perceptual processing of irrelevant items occurs, depends on the demands placed by the relevant task, with early selection of targets occurring under conditions of low perceptual load (thus no semantic processing of distractors), and late selection of targets occurring under conditions of high perceptual load (i.e., at the response level).

In light of the perceptual load theory of attention, the findings of Jiang and Leung (2005) and Cock et al. (2002) are more likely to have arisen as a consequence of attention spilling over to the to-be-ignored distractors under conditions of low perceptual load, and not to genuinely 'unattended' processing or learning. In particular, Jiang and Leung's study supports the idea that under conditions of low load, selection occurs only at the response stage of processing: they found no effect of the distractor set during the learning phase of the experiment; yet latent learning of this distractor set was 'released' when it subsequently became the focus of attention for the target task. In concordance with perceptual load theory, it should be possible to abolish learning of the to-be-ignored distractor sets in contextual cueing by increasing the perceptual load imposed by the attended items. In fact, an earlier study by Jiang and Chun (2001) found just that: by increasing the similarity between the target and the attended distractor set, Jiang and Chun (Experiment 4) discovered that learning of the ignored distractor set, which had been observed in earlier experiments, no longer occurred.

In line with the argument presented above, it is necessary to determine whether learning in the SRT task is modulated by the perceptual load of the primary task. If it were found that sequence learning under conditions of high perceptual load was abolished, then this would constitute further evidence against the claim that implicit sequence learning occurs automatically. An ideal method for investigating the effects of perceptual load on sequence learning is to examine the effects of irrelevant distractors on multiple sequence learning. Evidence of learning multiple sets of contingencies, where the secondary (irrelevant) set is acquired as an incidental side-effect of performing the primary (relevant) task, might be considered a powerful demonstration of the automatic nature of implicit learning. This would especially be so if it were shown that the secondary set was learned without attention.

An early study of this nature by Mayr (1996) showed that both a sequence of object identities (the primary sequence) and a sequence of locations could be learned simultaneously, and without cost to learning either of them alone. Mayr found that participants were faster to respond to an object when both its identity and the location at which it appeared were predictable compared with only one dimension being predictable. The conclusions drawn from that study – that independent systems exist for learning object and spatial sequences, that they operate without interfering with each other, and that attention is not necessary for learning the separate sequences - are quite convincing, given that unlike the Jiménez and Méndez (1999) study and the Cock at al. (2002) study, the independent sequences were not correlated. However, as explained in Chapter 1 (see also Shanks et al., 2005), the absence of a suitable single-sequence control group in Mayr's study obfuscated the evidence that two sequences can be learned independently of attention. In Mayr's experiments, the single sequence groups responded, for example, to an object sequence, and simultaneously ignored a random location sequence (and vice versa). Under these conditions it is likely that participants devote attentional resources to attempting to learn the random sequence and thus it cannot be claimed that attention was fully devoted to learning the target sequence. Ideally, Mayr should have included single sequence groups where one dimension did not vary while the other did.

Yet a further need for caution when considering Mayr's claim that the independent sequences were learned without requiring attention, is that neuropsychological data has pointed to the existence of two different attentional systems: the anterior system, involved in target detection, and the posterior system,

involved in spatial orienting (Posner, 1992; Posner & Peterson, 1990). If it is the case that multiple pools of resources are available for separate attentional functions (e.g., Duncan et al., 1997; Navon & Gopher, 1979), then it is not at all clear whether Mayr's study gets to the heart of the matter about whether attention is necessary for implicit learning.

In light of these criticisms, the notion of an attentionally-independent implicit learning system capable of learning multiple contingencies has not been demonstrated beyond doubt. What needs to be known is whether it is possible to learn about multiple uncorrelated contingencies within the same domain (e.g., both spatial), *and* to what extent this depends on attention. The experiments reported in the present chapter were designed to provide new evidence on this issue. They are also reported in Rowland and Shanks (2006b).

Before describing the experiments it is important to clarify the way that attention is characterised for the present purposes. In line with the perceptual load theory (Lavie, 1995, 2005) of attention described above, the notion of resources and selection are not treated as independent. Thus, it will not be possible from the evidence presented in this chapter to distinguish between the relative contributions to learning made by these two forms of attention. This is not problematic, because the main goal of this chapter is to demonstrate how increasing the processing demands of the relevant (primary) task modulates learning. By exhausting the availability of attentional capacity for the irrelevant (secondary task), it reduces the opportunity for paying selective attention to that task. But it will not be clear whether this manipulation had the effect of limiting the amount of processing resources available for attention to be spread in parallel over the stimuli (thus not allowing early processing of irrelevant items), or whether it had the effect of preventing shifts of selective attention to irrelevant items in a serial fashion once relevant processing was complete. Nevertheless, the experiments are capable of addressing the pertinent issue of whether multiple sequence learning is abolished under conditions that demand that more attention be paid to the primary task.

It is expected that learning of multiple independent contingencies in a sequence learning task will occur provided the primary task, as a consequence of being undemanding, allows spill-over of attention to a secondary set of contingencies (investigated in Experiments 8 & 9). However, learning of a secondary sequence is unlikely if the demands of responding to the primary sequence are so high as to preclude the spill-over of attention (investigated in Experiment 9).

## 5.1 Experiment 8

This experiment was designed to determine whether participants are able to simultaneously learn two independent probabilistic sequences, one that is the focus of the experimental task and another that is secondary. Similar to Experiment 7 in Chapter 4, participants responded to a primary sequence during training whilst ignoring a secondary sequence, and then during the test stages they responded to the primary, the secondary, and a new sequence. One purpose of the investigation was to determine whether participants learned more about the secondary sequence than the new sequence. This would show that learning of a secondary independent sequence is possible. It was also predicted that the primary sequence would be learned most fully because it is the focus of attention in the experimental task.

There were two important modifications to the design in this experiment that sought to overcome some of the limitations of Experiment 7. First, training was extended considerably to allow a far greater degree of learning of the primary sequence, and, if it occurs, learning of the secondary sequence. It seems plausible that if secondary sequence learning is possible then it will either take longer to develop than primary sequence learning (because it is not the focus of attention), or it will occur only once the primary sequence has been learned fully at which point attention can be devoted to learning about less relevant contingencies. The second modification was to adopt a within-participants design. There is huge variation in the ability of people to learn these sequences, and thus a within-participants design will provide a far more powerful test of the hypothesis that the learning system can acquire multiple contingencies simultaneously.

### 5.1.1 Method

#### *5.1.1.1 Participants*

Twelve healthy volunteers participated in this experiment (5 women and 7 men, mean age = 22.3 years, range = 19-31). They were told that the purpose of the study was to see how fast they could become at responding to a repeatedly presented stimulus.

#### *5.1.1.2 Procedure*

Participants performed 48 blocks of 100 trials each of the SRT task, which comprised 6 blocks of training followed by 6 blocks of testing, repeated 4 times. Thus

there were 24 training blocks and 24 testing blocks. During the training blocks participants reacted to the location of a red circle in the bottom row (the target stimulus) and were instructed to ignore the red circle in the top row. Both stimuli disappeared immediately upon responding (i.e., the RSI = 0 milliseconds), and appeared at their respective sequence locations immediately. Their offsets and onsets were synchronised. Figure 5.1 panel **a** shows the display, which was identical to that of Experiment 7.

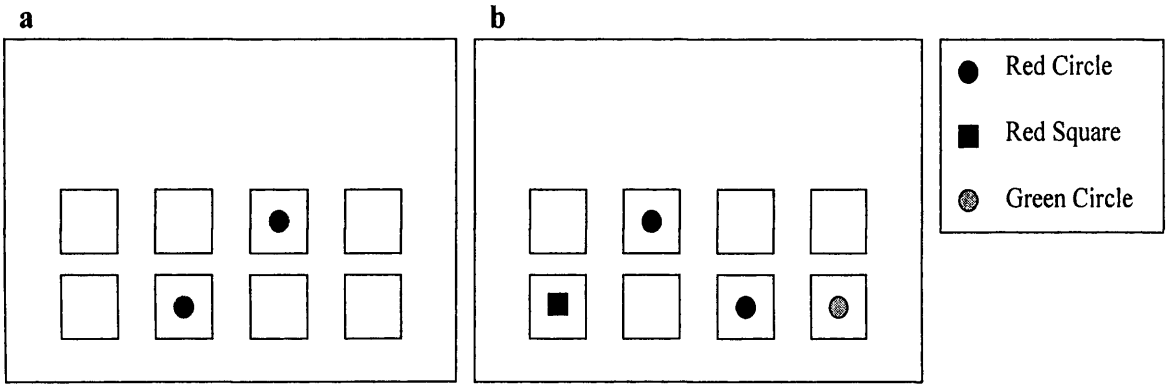


Figure 5.1 The displays for Experiments 8 (panel a) and 9 (panels a and b). The SRT display had a bottom and a top row of stimulus locations. On the training blocks participants responded to the target (a red circle) on the bottom row and ignored the top row. On the test blocks they responded to the red circle in the top row and ignored the bottom row. In Experiment 8 and for the low load group in Experiment 9 the display was as shown in panel **a**, whereas for the high load group in Experiment 9 the display was as shown in panel **b**. Participants in this latter group responded to the red circle and ignored the distractors. During the test blocks the distractors were removed.

For the first 2 blocks of the testing phase the target sequence was presented in the bottom row, as above. During the remaining 4 blocks of the test sessions participants were required to respond to the stimulus in the top row and ignore the stimulus in the bottom row (which continued to follow the attended sequence). For two successive blocks of responding to the target in the top row the sequence was the secondary sequence (that is, the ignored sequence that appeared in the top row during the training sessions), and for the other two was a new sequence. This new sequence was identical for all 'new' testing blocks. The order of the testing blocks for secondary and new sequences was counterbalanced. In short, the experiment examined learning of a primary, a secondary, and a new sequence over a 48 block SRT session.

*5.1.1.3 Sequence Information*

Six sequences arranged into 3 pairs were used: pair 1 (2-1-2-3-4-1-3-2-4-3-1-4 and 4-1-4-3-2-1-3-4-2-3-1-2); pair 2 (3-1-3-2-4-1-2-3-4-2-1-4 and 4-1-4-2-3-1-2-4-3-2-1-3); and pair 3 (1-2-1-3-4-2-3-1-4-3-2-4 and 4-2-4-3-1-2-3-4-1-3-2-1). One sequence of the pair generated the probable trials, and the other the improbable, or 'noisy', trials. The probability of a target location trial being consistent with the primary sequence was .85 and the probability of a trial being 'noisy' was .15. Assignment of sequence pairs to conditions, and the selection of the probable and noisy sequence within a pair, was determined randomly. Unavoidably, the sequences do share some transitions. For example, sequence 1 from pair 1 and sequence 1 from pair 2 both contain the transition 1-2-3. However, by randomly assigning the sequences to conditions and to probable and improbable trials any imbalance or regularity of transition similarities between conditions is avoided as any such effects should be cancelled out. This method of sequence generation ensured that the primary and secondary sequences were statistically independent, as briefly explained below.

Consider the sequence pair Probable-1 (2-1-2-3-4-1-3-2-4-3-1-4) and Improbable-1 (4-1-4-3-2-1-3-4-2-3-1-2). If a sequence transition such as 2-1 (from Probable-1) occurs, the probability of the next location in the sequence being 2 is .85, and of being 3 (which is the location following 2-1 in Improbable-1) is .15. This method iterated on each trial, such that 2 typical location sequences might therefore be (a) 21***3***243***2***4312***3*** and (b) 2123***1***421***3***243 where the highlighted symbols refer to improbable locations. Because the introduction of improbable trials is unpredictable for each sequence, this method of sequence generation ensured that there was no location correlation between sequences: for example, in (a) the first occurrence of the triplet ***3***24 co-occurs with the triplet 23***1*** in (b) whereas the second occurrence co-occurs with the triplet 421.

5.1.2 Results

*5.1.2.1 Training Phase Data*

Mean reaction times were computed by averaging across all correct trials separately for probable and improbable targets. This was done separately for each sequence type and for training and test sessions. Figure 5.2 shows the RTs to probable and improbable trials across the training blocks of the SRT task. From Block 6 onwards a healthy learning effect is apparent, suggesting that participants learned the primary

sequence. An ANOVA with Target Probability and Block as within-participants factors confirmed this by a main effect of Target Probability, $F(1, 11) = 78.09$, $MSE = 1,377$, $p < .01$. Predictably, learning improved with practice, as shown by the Block × Probability interaction, $F(31, 341) = 1.54$, $MSE = 651$, $p < .05$.

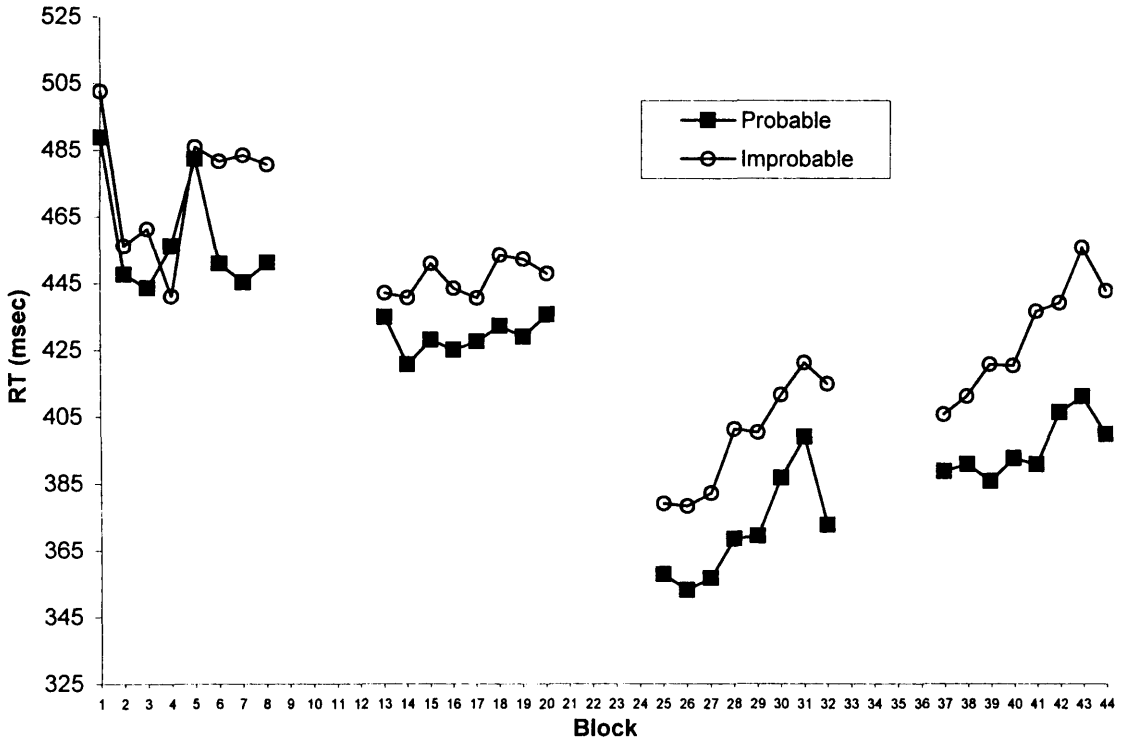


Figure 5.2 Mean RTs to probable and improbable targets across the training blocks (responding to the primary sequence in the bottom row) of Experiment 8. The gaps in the plots signify the points where participants switched to responding to the sequence in the top row.

*5.1.2.2 Test Phase Data*

Figure 5.3 shows that learning scores for the primary sequence during the test blocks are higher than those for the secondary sequence and the new sequence; yet, as testing progresses, the learning scores for the secondary sequence are also considerably higher than those for the new sequence (which barely deviate from zero).

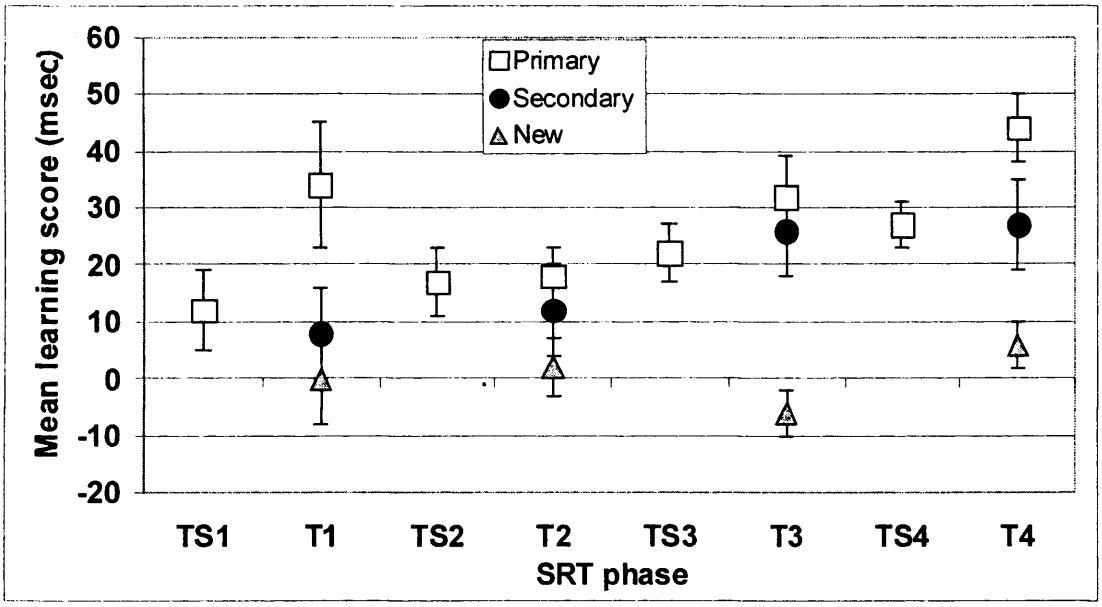


Figure 5.3 Mean learning scores (improbable RTs minus probable RTs) for the training sessions (TS) and the test sessions (T) of the SRT task on Experiment 8. Each training session comprised 6 blocks and was performed exclusively on the primary sequence. Each test session comprised 6 blocks, 2 on the primary sequence, 2 on the secondary sequence and 2 on the new sequence (the order of the secondary and new sequences was counterbalanced). Error bars depict standard errors of the means.

The first analysis on the test blocks compared the effects of target probability between the new and secondary sequences across the 4 test sessions. A Sequence (new vs. secondary) × Target Probability ANOVA revealed no main effect of sequence, $F(2, 22) = 2.58$, $p = .10$, $MSE = 304$, but the analysis did confirm a main effect of target probability, $F(1, 11) = 54.23$, $p < .01$, and most importantly, a Sequence × Probability interaction, $F(1, 11) = 30.82$, $MSE = 583$, $p < .01$, thus showing that participants learned the secondary sequence during the test sessions but did not learn the new sequence. The Block × Sequence × Probability interaction also reached significance, $F(5, 50) = 2.53$, $MSE = 710$, which is reflected in the secondary versus new comparisons on the first 3 test sessions (session 4 is dealt with separately in the next paragraph): there was a secondary-new difference at test session 3, $t(11) = 2.54$, $p < .05$, two-tailed, but not at test sessions 1 or 2 (*n.s.*). The order in which participants encountered the new and secondary sequences (which was counterbalanced) during the test sessions was also included as a factor in the analyses. None of the statistics that included order as a factor reached significance (all $F$s < 1.67).

The critical comparisons were on the learning scores for each sequence on the final test sessions, as differences in learning were expected to be most apparent at the end of training. A one-way ANOVA on the learning scores for each condition on the final test blocks revealed a main effect of sequence type, $F(2, 22) = 15.10$, $MSE = 326$, $p < .01$. The primary sequence had a learning score greater than that of the secondary sequence, $t(11) = 1.87$, $p = .045$, one-tailed, and in turn, the secondary sequence had a learning score greater than that of the new sequence, $t(11) = 3.79$, $p = .003$, two-tailed. Only the primary and secondary sequences had learning scores that were greater than zero, $t(11) = 6.97$, and $t(11) = 7.92$, $ps < .01$, two-tailed, respectively.

In general, test-phase learning scores for the primary sequence were greater than those on the block immediately after the test blocks, by an average of 12 and 8 msec in Experiments 8 and 9 (see section 5.4.3), respectively. This suggests that the primary sequence learning scores included in the analyses were slightly inflated due to the fact that the primary sequence was always tested immediately after training blocks with the same sequence. However, separate ANOVAs for each experiment on these data, with pre-switch/post-switch and training/test cycle as the factors (the design precluded this analysis at test phase 4), revealed no reliable effect of position (pre- vs. post-switch) nor any interaction with cycle (all $ps > .05$). Thus, despite a slight average reduction in the magnitude of learning scores upon returning to the primary sequence after switching, the effect was not significant.

To take account of the differences in absolute RTs when switching between sequences (<50 msec in all cases) proportional measures were computed [(improbable RT - probable RT)/probable RT] for each condition on the final test blocks. The obtained values, .11, .06, and .02, (primary, secondary, and new, respectively) mirrored the pattern found for the learning scores, and differed statistically, $F(2, 22) = 16.67$, $MSE = .002$, $p < .01$. Crucially, the secondary vs. new difference was also significant, $t(11) = 3.99$, $p < .01$, two-tailed. It is unlikely therefore that a scaling effect is responsible for the differences in learning scores.

### *5.1.3* Discussion

The results from this experiment are very clear: participants learned more about a secondary sequence than they did about a new sequence, but not to the same degree as they learned about a primary sequence. This demonstrates learning of multiple independent sequences.

## 5.2 Experiment 9

Experiment 8 suggests that learning about a secondary sequence can take place without attention. However, an alternative possibility is that during the training stages of the SRT task, particularly in later stages, responding to the primary sequence did not fully engage attention because the task had become easier to perform through practice. This allowed unused attention to spill over to learning about the secondary sequence. To test this conjecture, the following experiment increases the perceptual load of the primary task by adding distractor items to the attended display locations. This should result in increased attention to the primary sequence, and therefore a reduction in the amount of attention available for learning about the secondary sequence. It is hypothesised that this manipulation will reduce the amount of learning about the secondary sequence, but not that of the primary sequence (which still engages full attention).

### 5.2.1 Method

#### *5.2.1.1 Participants*

A further 32 volunteers took part in this experiment (17 women and 15 men, mean age = 23.0 years, range 20-42). Half were assigned to a low load group and half to a high load group.

#### *5.2.1.2 Procedure*

For the low load group the procedure was exactly the same as for Experiment 8. For the high load group the experiment was exactly the same as Experiment 8 but the bottom row of the SRT display contained two additional stimuli on each trial: a filled red square and a filled green circle. Note that each of these shares one feature with the target. This ensured that participants in the high load group would have to pay more attention to selecting the target due to competition from the distractors (Treisman et al., 1983). The distractor stimuli appeared pseudo-randomly: on any trial they never shared a location either with each other or the target, and they never appeared at immediately successive locations. Figure 5.1 panel **b** shows the display for the high load group.

To ensure the maximal recruitment of attention to the primary sequence throughout training in the high load group the distractors were present on all blocks of responding to the primary sequence except for the last two (i.e., Blocks 43 & 44).

Removing the distractors for these blocks allowed for a measure of learning under low load conditions for the primary sequence. To ensure parity of conditions with the low load group, all blocks of responding to the secondary and new sequences in the upper row were performed without the presence of the distractors. Thus, only the final 6 blocks (2 primary, 2 secondary, 2 new) allowed a comparison with identical stimuli between the 2 groups for all 3 sequences.

### 5.2.2 Results

#### *5.2.2.1 Training Phase Data*

Figure 5.4 shows the RT data for the training stages of the SRT task. With the exception of Blocks 43 and 44, RTs for the high load group on the training blocks were between 100-200 milliseconds slower than those for the low load group. When the distractors were removed for Blocks 43 and 44, RTs were roughly equivalent to those of the low load group. This shows that the distractors had a sustained performance effect by continually forcing participants to attend to the target sequence.

The RT data for the training blocks of the primary sequence were entered into an ANOVA with Target Probability and Block as within-participants factors, and Group as a between-participants factor. As predicted, both groups learned the primary sequence, as shown by the main effect of target probability, $F(1, 30) = 47.75$, $MSE = 3{,}358$, $p < .01$, although, with practice, this effect became greater for the low load group as confirmed by a Probability × Group × Block interaction, $F(29, 870) = 1.50$, $MSE = 1{,}532$, $p < .05$.

#### *5.2.2.2 Test Phase Data*

The crucial data were the groups' learning scores for each of the sequences on the final blocks, where testing for both the low and high load groups occurred under identical conditions. They are shown in Figure 5.5. Clearly, learning scores for the secondary sequence are lower for the high load than for the low load group, whereas this is not the case for the primary and the new sequences. Indeed, this is confirmed by a Sequence × Group interaction, $F(2,60) = 3.50$, $MSE = 611$, $p < .05$, and independent t-tests which show that learning scores differed for the secondary sequence, $t(30) = 2.39$, $p < .05$, two-tailed, but not for the other sequences (*n.s.*). For the primary sequence both groups obtained learning scores greater than zero ($ts > 5.0$, $ps < .01$), as did the low load

group for the secondary sequence, $t(15) = 4.88$, $p < .01$. No other learning scores deviated from zero.

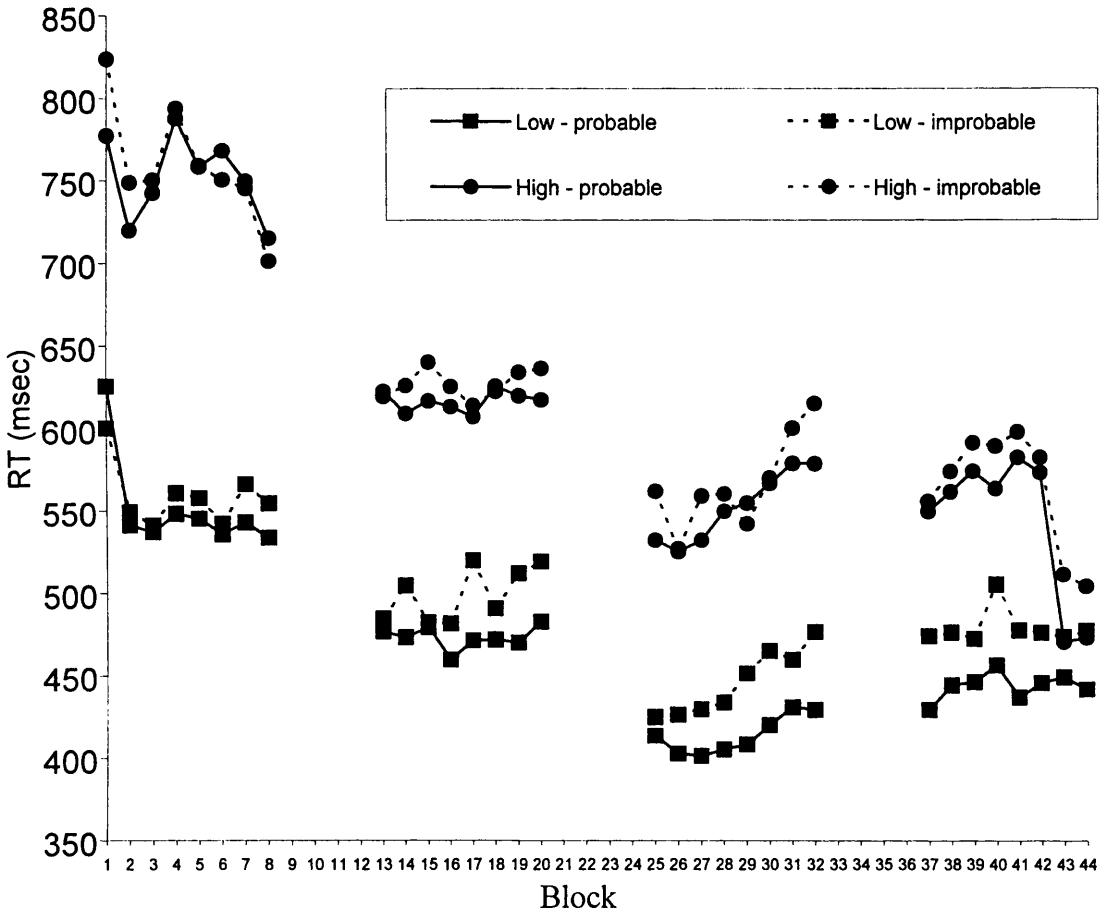


Figure 5.4 This graph plots the RTs to probable and improbable trials for the primary sequence for both the low and high load groups of Experiment 9. The gaps represent the point where participants responded to the sequences in the top row. Clearly, performance in the high load group is impaired compared with that of the low load group, whose RTs are on average over 100msecs faster than the high load group. The sudden increase in RT speed on blocks 43 and 44 for the high load group is due to the removal of the distractors.

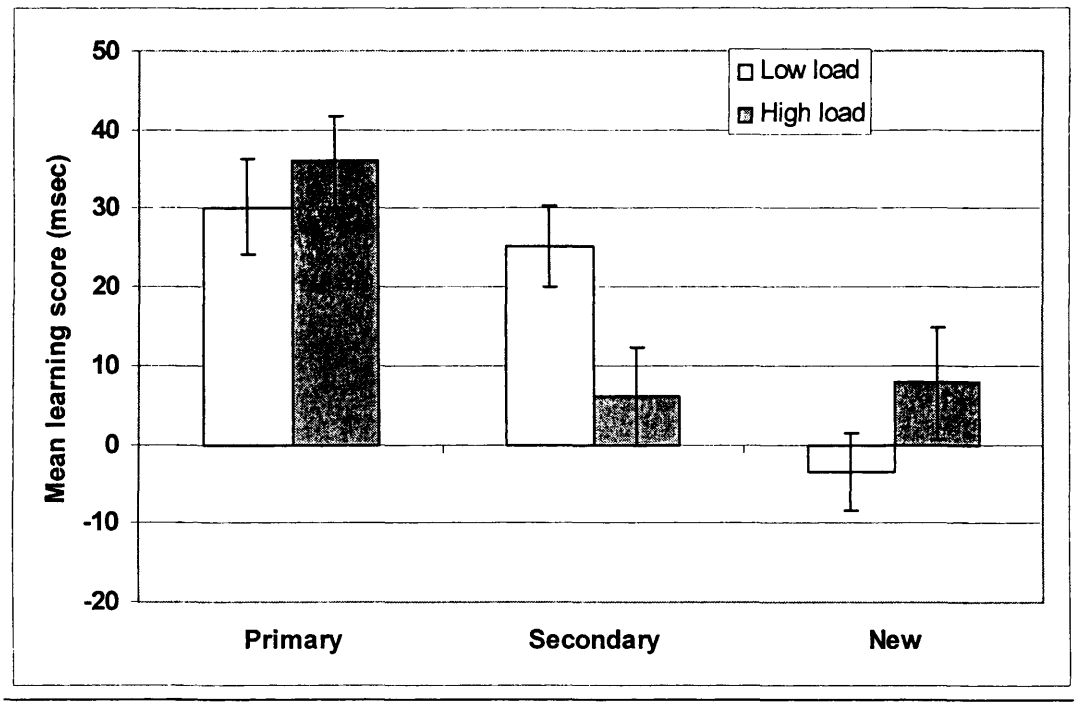


Figure 5.5 Mean learning scores (improbable RTs minus probable RTs) for each sequence type on the final test blocks for the low load and high load groups of Experiment 9. Although learning scores for both groups are very similar for the primary sequence, the low load group has a markedly lower learning score than does the high load group for the secondary sequence. Error bars depict standard errors of the means.

As in Experiment 8, proportional measures mirrored the pattern of the learning scores: .07, low/primary; .08, high/primary; .06, low/secondary; .01, high/secondary; -.01, low/new; .02, high/new. These measures differed statistically, as shown by a reliable Sequence × Group interaction, $F(2, 60) = 3.22$, $MSE = .003$, $p < .05$, and the low and high groups had statistically different proportional measures for the secondary sequence, $t(30) = 2.54$, $p < .05$.

### 5.2.3 Discussion

These results replicate the secondary-new difference of Experiment 8 and confirm, firstly, the learning of multiple independent sequences, and secondly, that this learning is attenuated by an increase in the perceptual load of the primary task. Although it is the secondary-new difference that is of principal interest, the primary-secondary difference is also worth comment. In contrast to Experiment 8, the difference in learning scores between the primary and secondary sequences for the low load group

was not statistically significant. But in both experiments, learning scores increased little over training after the first training session for the primary sequence, and from previous work with these sequences it is known that learning scores under single-task conditions peak at approximately 40 milliseconds (Shanks et al., 2005). It is thus likely that a ceiling effect is responsible for the finding that learning for the primary sequence is no greater than that for the secondary sequence. Thus it would be unwise to place undue weight on the rather small differences observed between the primary and secondary sequences.

## 5.3 Discussion of Experiments 8 and 9

The effect on sequence learning of increased perceptual load is consistent with Lavie's (e.g., 2005) theory of attention. In her experiments the allocation of attention is highly dependent on the degree and type of load under which participants perform. Under conditions of low perceptual load of the primary task, available resources not consumed by the task (because selecting the targets is easy) allow for the spill-over of selective attention to processing task-irrelevant items. However, if the perceptual load of the primary task is high then selecting the relevant target(s) is more difficult. Therefore, to ensure efficient processing, more resources are required to control the locus of selective attention, and thus fewer resources remain for attention to spill-over to irrelevant items. It should be stressed that reductions in distractor processing under high perceptual load cannot be explained as a result of increased task difficulty, decreased processing speed, or by additional stimuli impinging on sensory processing limits (or, data limits, Norman & Bobrow, 1975), but instead, as a result of a reduction in available attentional resources (Lavie & de Fockert, 2003). This accounts for the abolishment of secondary sequence learning under conditions of high perceptual load. The present experiments thus lead one to surmise that the results obtained by Jiang and Leung (2005) and Cock et al. (2002) were due, not to pure attention-independent learning, but because spare processing resources were available for learning about supposedly ignored contingencies. Overall, the present findings are consistent with the hypothesis that implicit learning does require attention.

Lavie's theory distinguishes between different types of attentional load. In contrast to the effects of perceptual load, high working memory (or cognitive) load appears to increase interference from irrelevant distractors because working memory load reduces the capacity of control functions. This explains why primary sequence

learning in Experiment 9 (and the experiments in Chapter 3) of the present study was unaffected by an increase in perceptual load (see also Rowland & Shanks, 2006a, 2006b), whereas primary sequence learning in Experiment 1 was attenuated by an increase in working memory load. Under the former conditions, adequate control ensures reduced interference from distractors, whereas under the latter conditions, a reduction in control allows for greater distraction, and thus reduced sequence learning. The finding that multiple sequences can be learned simultaneously, and that primary sequence learning is unaffected by increased perceptual load, suggests that implicit learning, unlike explicit learning, is highly resistant to noise, or input complexity. Conversely, like explicit learning, reduced control due to working memory load leads to a reduction in implicit learning.

It may be argued (Luis Jiménez, personal communication, July 2005) that the inclusion of random distractor items in the SRT task makes the display noisier and thus participants find it more difficult to segregate the useful dimensions from the irrelevant ones. Perhaps this, instead of reduced attention, accounts for the decrease in learning about the secondary sequence? There are several reasons why this may not be the case. First, between blocks 9-12 participants responded to the secondary sequence and thus it becomes clear early in the task which is the useful distractor dimension to which to attend. Second, because the primary distractors and the secondary sequence were in separate rows, it is clear that there is a distinction between them. It seems improbable that participants would attempt to aggregate information between rows rather than within rows. And finally, following the preceding arguments, an increase in noise due to the inclusion of distractors would likely impair learning of the primary sequence either in tandem, or instead of, the secondary sequence. There was no evidence of this here. Therefore an attentional explanation for the findings is favoured.

A remarkable feature of the present study and the other studies upon which this chapter has focussed is that complex information about the secondary sequence is gathered in a seemingly passive fashion, even though the secondary sequence is, on the surface, irrelevant to the primary task. These findings are testament to the sophistication of the human learning system, but should not be taken as evidence for a mysterious mental faculty. Instead, this passive perceptual learning appears to be dependent on the attention-grabbing features of the secondary contingency set. In the present experiments and those of Cock et al. (2002), the secondary sequence was borne by a separate visual stimulus from that of the primary sequence. This was either at a unique set of locations

(in the present experiments) or was a different colour from the primary target (in Cock et al.'s experiments). In the Mayr (1996) and Jiménez and Méndez (1999) studies, both dimensions were presented on the same visual target. However, Mayr ensured a large separation between locations for the secondary spatial sequence, and Jiménez and Méndez required that the targets for the secondary sequence are classified – importantly, secondary sequence learning disappeared when classification was unnecessary. In each of these cases the secondary dimension vied for attention with the response dimension. Conceivably, multiple contingency learning is dependent on the availability of resources allowing selective visual attention to be shared amongst the attention-demanding tasks.

As the foregoing point suggests, the development of secondary sequence knowledge may be largely dependent on attention being drawn to the upper row of distractors because, once they have been seen to be relevant to the experiment (during testing), they then become an important dimension to which to attend. An interesting question is whether evidence of secondary sequence learning would have been obtained had testing on this sequence only taken place at the very end of training (i.e., if there had not been intermittent testing bocks)? The results from Experiment 7 suggest that this would not be the case, but it is essential to provide the same amount of training as that provided in Experiments 8 and 9, and to employ within-participants designs. This is a very interesting line of enquiry, and one that would be worth exploring further. If secondary sequence learning were obtained without attention being 'drawn' to the irrelevant dimension, this would be a powerful demonstration indeed of the ability of an implicit learning system to track and learn about seemingly unattended statistical regularities.

In summary, this chapter has demonstrated a critical role for attention in learning multiple contingencies in a probabilistic sequence learning task. Future work could help to determine: (i) exactly what is learned about the secondary sequence and to what extent this is implicit or explicit; (ii) whether learning of the secondary sequence occurs contemporaneously with learning about the primary sequence, or develops only once the primary sequence is automatised; and, (iii) perhaps by eye-tracking, whether attention to the secondary sequence is covert or overt.

# Chapter 6

## What Remains of the Idea that Implicit Learning is an Automatic Process?

The principal objective of the present work was to evaluate the proposition that implicit learning is a unique form of learning that can be distinguished by its ability to operate automatically. To that end, this final chapter re-evaluates the notion of automatic implicit learning in light of the evidence contributed by the experiments reported in this thesis.

### 6.1 Does Implicit Sequence Learning Require Attentional Resources?

Experiment 1 in Chapter 2 investigated probabilistic sequence learning in the presence of a symbol-counting secondary task. The idea behind the dual-task SRT procedure is that the localisation task competes with the secondary task for mental resources and thus reduces the amount of attention available for sequence learning. The attenuation of learning under these conditions has been taken as evidence that implicit sequence learning is not, in this respect, an automatic process.

The finding in Experiment 1 was that symbol counting did impair probabilistic sequence learning relative to a single-task group trained without the secondary task. This result is consistent with studies that observed decrements in learning under dual-task conditions (e.g., Nissen & Bullemer, 1987; Shanks & Channon, 2002) and with work that used probabilistic sequences and a symbol-counting secondary task (Shanks et al., 2005). However, this observation is at variance with many other studies that did not obtain a reduction in learning under dual-task conditions (e.g., Frensch et al., 1998; Jiménez and Méndez , 1999, 2001; Schvaneveldt & Gomez, 1998).

Many of the earlier investigations of dual-task sequence learning were plagued by the methodological problems of using deterministic sequences and tone-counting secondary tasks, and it has thus been difficult to draw firm conclusions from these studies collectively (Shanks, 2003). The present Experiment 1 and the study by Jiménez and Méndez (1999) sought to avoid these problems, and although they share similar design features, the findings from the two studies conflict with each other. As argued in Chapter 2, it seems that the advantages of very extensive training, a 240-millisecond response stimulus interval, and an additional predictive relationship between symbols and locations in the Jiménez and Méndez study may have combined to offset the demands of the secondary task during learning. Although in their paper Jiménez and

Méndez dispute this explanation of their findings (see *General Discussion*, Jiménez & Méndez, 1999), the concerns about the effects of the conditions they created resurface when considered with respect to the present Experiment 1, which, by minimising the possible benefits afforded by the above features, found an impairment in dual-task learning. Given that this result is consistent with a vast amount of work that supports the resource-dependent hypothesis of implicit learning (e.g., Hsiao & Reber, 1998; Perruchet & Vinter, 2002), implicit memory (e.g., Mulligan & Brown, 2003), human classical conditioning (e.g., Carter, Hofstötter, Tsuchiya, & Koch, 2003), and automatic processing more generally (e.g., Logan et al., 1999), the more parsimonious theoretical account is that implicit learning is not unique in the demands it places on attentional resources.

However, it would be informative to refine this statement somewhat. To be precise, Experiment 1 shows that a secondary symbol-counting task attenuates probabilistic sequence learning. The symbol-counting task undoubtedly depends on working memory for continuously updating the total number of relevant symbols seen. It is more accurate then to state that implicit sequence learning is dependent on the availability of working memory. This approach is potentially more meaningful for several reasons. First, there is controversy about the concept of resources (Navon, 1984; Neisser, 1976), making it very difficult to empirically ascertain to what extent attentional resources are involved in processing. Thus, it is always possible that target processing under dual-task conditions remains unimpaired not because it does not consume resources, but because capacity has not been exhausted (indeed, Frensch and Miner, 1994, demonstrated that implicit learning is only affected when capacity limits are reached). Second, the relationship between resources and other types of attention is far from clear. Lavie (2005) has shown that selection processes and resource operations are closely intertwined, and thus it might not be advisable to treat them as independent. Finally, it is not clear exactly how secondary tasks have their effects (Pashler, et al., 2001; Rah et al., 2000), and thus it might be unwise to claim that learning is independent of resources, when all that has been shown is that learning is not affected by the secondary task.

In contrast, one can be certain that the counting task does rely on some form of working memory, and thus it is accurate to claim that sequence learning is impaired by a reduction in working memory capacity. Working memory is itself considered a form of attentional operation (Baddeley, 1996; de Fockert et al., 2001; Humphreys &

Samson, 2004) that is neurally almost indistinguishable from attention networks in the brain (Pessoa & Ungerleider, 2004). In this sense, working memory is a control process that carries out many functions such as holding information in mind, configuring task sets, directing selective attention and action, and directing attention towards information relevant to current behavioural goals (Baddeley, 1996; Engle, 2002). The inclusion of a secondary task taxes working memory by increasing both the number of activated representations that it must deal with (in the symbol-counting task, spatial location, symbol identity, and the serial position of at least the current item and the previous two, will be simultaneously held on-line) and the complexity of control processes needed to keep in mind which symbols are to be counted, which are to be ignored, and to refresh and retain the count. With so many competing events, the effective capacity of working memory is reduced, and it becomes more difficult for earlier targets to be maintained in position so that they can be associated with subsequent targets (Hsiao & Reber, 2001).

The conclusion from these studies - that implicit learning is dependent on working memory - is thus consistent with the notion that learning generally requires working memory to associate relevant events in short-term memory, and to ensure that attention is directed appropriately to relevant information.

### 6.2 Does Implicit Learning Require Selective Attention?

Automatic processes may also be distinguished by their independence of selectional control (Schneider & Shiffrin, 1977). It has been suggested however that implicit learning is dependent on selective attention, and the results of Experiments 6-9 presented in Chapters 4 and 5 of the present thesis are consistent with this proposal. In these experiments, a target sequence and a secondary (to-be-ignored) sequence were presented simultaneously. In Experiments 6 and 7 it was found that learning about the secondary sequence did not take place – although there was some rather weak evidence from Experiment 6 showing that the secondary sequence had been learned, but this learning was accompanied by strong negative priming effects for the secondary sequence, which are generally considered to arise from attentional processes (Houghton & Tipper, 1994; Lavie & Fox, 2000).

Experiments 8 and 9 offered a new approach to studying the role of attention in sequence learning. Experiment 8 established that a secondary spatial sequence could be learned in tandem with a primary spatial sequence with which it was not correlated. Previously it had not been shown that multiple statistically independent probabilisitc

spatial sequences can be learned simultaneously. This is an interesting finding because it is a powerful demonstration of the sophistication of the human implicit learning system, which corroborates claims that implicit learning is accurately encapsulated as a process that operates incidentally (Jiménez, 2003) and that is able to detect multiple patterns of covariation (Cleeremans, 1993; Reber, 1993). However, Experiment 9 showed that there are limits on the capabilities of such learning: the secondary sequence learning observed in Experiment 8 was abolished under conditions that restricted the availability of attention to spill over to learning about the secondary set of contingencies.

The finding that increasing the task difficulty of a primary sequence attenuates learning of a secondary sequence is consistent with theories of attention that posit that the locus of selective attention is modulated by attentional load (Lavie, 1995). Additionally, the notion that attention during encoding is a necessary condition for learning (Logan et al., 1999) lends credence to the proposal herein that selectively attending to sequence events is critical for associating them during the course of implicit learning. Moreover, this view is consistent with the idea expounded above that implicit learning is dependent on the availability of working memory. Working memory is essential for the controlled and successful deployment of selective attention (de Fockert et al., 2001) and for elementary associative processes (Carter et al., 2003). From this perspective, implicit learning cannot be considered an automatic process.

## 6.3 Multiple Sequence Learning

An interesting case in support of the claim that implicit learning is automatic is found in studies that demonstrate multiple sequence learning. Presumably, if a secondary sequence that is uncorrelated with the primary sequence – and is incidental to the primary task – is learned, and learning this secondary sequence does not impair learning about the primary sequence, is this not strong evidence for automatic learning? In one such study, Mayr (1996) reported learning of both an object sequence and a spatial sequence that were uncorrelated with each other, and that learning two sequences imposed no costs on learning a single sequence. However, as discussed in Chapter 1, Mayr did not include an appropriate control group that would have fully evaluated this latter result. Moreover, as Jiménez and Méndez (1999) highlighted, due to the large spatial separation of the locations, participants were likely to have selectively attended to them (in this sense, actively oriented to them) in the process of learning the spatial

sequence. It is not warranted then to consider Mayr's study as having supported the existence of automatic learning. But the fact that two uncorrelated sequences were learned simultaneously is in itself tantalising, especially as the secondary sequence was incidental to the primary sequence. However, contrary to the claims of automaticity, Experiment 9 showed multiple sequence learning to be a phenomenon that is dependent on the availability of attention.

## 6.4 Selection Costs and the Resilience of Implicit Learning

The preceding sections argued that it is not consistent with the available evidence to claim that implicit learning operates automatically, either with respect to demands on attentional resources or selective attention. Is there any sense in which implicit learning should be called 'implicit'? If it does not possess any unique properties, then should it not be simply termed 'learning'? This will be considered below, but first the results from Experiments 2-5 will be discussed because they do seem to show some interesting features of learning that are untypical.

Chapter 3 presents some experiments that interfere with task performance by introducing distractors in the SRT procedure. Following previous work on perceptual filtering (Kahneman et al., 1983; Treisman et al., 1983), it was expected that the inclusion of salient perceptual distractors would compete with the target for attention, and thus slow reaction times to the target stimuli. This 'filtering cost', or 'selection cost', did have marked impairments on SRT performance, slowing reaction times by up to 400 milliseconds. However, the interesting result was that despite the cost to performance, learning was unimpaired across five experiments and under a variety of conditions. It was concluded in that chapter that disrupting the target selection process does not, at least under the conditions presented here, impair learning. How is this result to be interpreted? At first glance it reflects a result that is consistent with the automatic proposal of implicit learning, in that implicit learning is resistant to divisions of attention during encoding. But this finding does not contradict the view that implicit learning is dependent on working memory. Instead, the interpretation adopted here is that implicit learning may be impervious to load on input (or perceptual) attention, and yet highly dependent on working memory. This distinction is consistent with some recent models of attention that delineate two capacity limited components of attention, an input/selective mechanism, and a central/control, mechanism (e.g., Johnston et al., 1995; Lavie et al., 2004). Experiments 2-5 impose a minimal (possibly a negligible)

working memory load, given that, on each trial, participants simply have to respond to one stimulus and ignore another on each trial. However, the performance decrements imposed by the noisy displays suggest that there was strong competition for attention at the perceptual-motor stage of the task. Yet no effect on learning was observed. The view taken here is that because working memory is available to cohere and segregate the inputs, the interference from the distractors is restricted to the input stages of the task. This is an intriguing result because it is generally regarded that sequence learning is susceptible to increases in noise during encoding (Schmidtke & Heuer, 1997) and that interfering with the associative process will disrupt learning about a sequence (Frensch & Miner, 1994; Hsiao & Reber, 2001). But in contrast to the experiments in this thesis, previous studies have investigated the effects of additional noise on sequence learning of deterministic sequences, and thus these studies may not speak to the issue of whether *implicit* learning is resistant to noise. Probabilistic sequence learning is disposed to engaging implicit mechanisms, and consequently, different encoding processes might be operating. Disruption of the selection process during explicit sequence learning is highly likely to severely interfere with intentional and conscious strategies aimed at remembering the order of sequence transitions (because these control strategies reduce working memory capacity), but such strategies are recruited less during implicit learning (Jiménez, 2003). During the slow accrual of sensitivity to the statistical properties of noisy sequences (Cleeremans, 1993), the greater availability of working memory is sufficient for minimising the disruption caused by the irrelevant distractors.

The findings from Experiments 8 and 9 also are consistent with the proposal that implicit learning is resistant to noise: learning two complex noisy sequences simultaneously is an excellent example of successful learning under very noisy conditions. Undoubtedly, further research is needed to expand upon these provisional results.

## 6.5 Theoretical Orientations

The conception of implicit learning that emerges from the work in this thesis is that it is a form of learning that occurs incidentally in the presence of complex stimulus inputs. This comports with other models of implicit learning that characterise implicit learning as the unintentional detection of patterns of covariance in the environment (Cleeremans, 1993; Reber, 1993). Additional support for this view is provided by Saffran, Newport, Aslin, Tunick, and Barrueco (1997), who showed that participants

who are engaged in a visual task can become sensitive to the statistical features of complex auditory information played in the background. Where the view presented in this thesis departs from many other claims about implicit learning, is that the above operations are not considered to take place in the absence of demands on attention – instead, working memory availability is necessary for learning to remain intact. Indeed, the finding that working memory is critical for implicit learning is consistent with recent reports that individual differences in working memory capacity are strongly related to sequence learning (Unsworth & Engle, 2005), and other processes deemed to be automatic, such as fear conditioning (Carter et al., 2003). However, under conditions of minimal working memory demands, spare processing capacity is able to spill over to other elements in an implicit learning situation. Thus, attending to task irrelevant stimuli is likely to take place passively (Johnston & Dark, 1986, distinguish between passive and active selection) as a by-product of goal-relevant processing. No active attempts to learn are necessary. But a system operating in this manner is unlikely to be capable of processing deep properties of stimuli. Instead, implicit learning is most likely confined to relatively simple stimulus features, and the statistical regularities between them. In order to detect these features, implicit learning is most prominently active in environments that present a large quantity of repetitive information. This notion of implicit learning taking place slowly and incrementally is captured by an information processing account of implicit learning (St. John & Shanks, 1997). Other approaches too have viewed implicit learning in this fashion (e.g., Cleeremans & Jiménez , 2002; Perruchet & Vinter, 2002). However, where the present account differs, particularly with respect to Cleeremans & Jiménez (2002) is that there is no postulation of purely automatic or unconscious processes. Instead, implicit learning is more aptly characterised as a process that operates on complex patterned information by becoming increasingly attuned to the inherent regularities of that information, first by consciously detecting "fragments" of the input (e.g., Perruchet & Vinter, 2002), and then by attending more strongly to these conscious contents of working memory (Baars, 2002). It may be informative to characterise this process as being a weak form of automaticity: although control processes are necessary to guide attention and cohere working memory inputs, there are no controlled attempts to learn, nor is the learner required to actively select (i.e., respond to) the inputs.

## 6.6 Future Directions

The experiments in this thesis point to some interesting directions that implicit learning research could take. The finding that selection costs do not impair learning need to be extended using methods that investigate different approaches to imposing filtering demands. For example, it would be useful to determine to what degree implicit learning is resistant to input complexity. One prediction is that filtering demands that draw on working memory capacity are likely to impair learning, such as the requirement to select the target from highly salient distractors that share semantic features with the target.

Another potentially worthwhile approach is to investigate the effects of distractors under working memory load. For instance, participants could be trained on a sequence learning task that includes increasing levels of selection costs. As the competition between targets and distractors increases, the availability of working memory becomes more crucial. Perhaps a symbol counting task under these conditions would reveal greater impairments to sequence learning as selection costs increase?

The claim that implicit learning is resistant to noisy input needs to be validated with respect to explicit versions of implicit learning paradigms. It is generally accepted that deterministic sequences, particularly those with just a few transitions, are likely to be learned by recruiting conscious intentional strategies, and thus do impose substantial attentional demands (Jiménez, 2003; Jiménez & Vázquez, 2005). Therefore, according to the theoretical ideas set out here, filtering costs should show detrimental effects on learning deterministic sequences.

Finally, a complete understanding of the role that attention plays in learning will only be achieved by further consideration of the mechanisms that support human attention. This is a major challenge because the literature on attention comprises one of the richest and most pervasive in psychology and neuroscience. This thesis has made a tentative step towards this goal.

References

Allport, D. A. (1992). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In D. E. Meyers & S. Kornblum (Eds.), *Attention and performance* (Vol. XIV, pp. 183-218). Cambridge, MA: MIT Press.

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences, 6,* 47-52.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology, 49A*, 5-28.

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113,* 518-540.

Batsell, W. R. (2000). Augmentation: Synergistic conditioning in taste-aversion learning. *Current Directions in Psychological Science, 9*, 164-168.

Beer, J. S., Shimamura, A. P. & Knight, R. T. (2004). Frontal lobe contributions to executive control of cognitive and social behavior. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences,* 3rd Edition (pp. 1091-1104), MIT Press: Cambridge, MA

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon Press.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*, 215–222.

Carter, R.M., Hofstötter, C., Tsuchiya, N., & Koch, C. (2003). Working memory and fear conditioning. *Proceedings of the National Academy of Sciences, 100*, 1399-1404.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28-71.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10*, 360-365.

Chun, M. M., & Jiang, Y. (2003). Contextual cueing: Reciprocal influences between attention and implicit learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 43-68). Amsterdam & Philadelphia: John Benjamins.

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science, 280,* 77-81.

Clark, R. E., Manns, J. R., and Squire, L. R. (2001). Trace and delay eyeblink conditioning: Contrasting phenomena of declarative and nondeclarative memory. *Psychological Science, 12,* 304-308.

Cleeremans, A. (1993). *Mechanisms of Implicit Learning: Connectionist Models of Sequence Processing.* Cambridge, MA: MIT Press.

Cleeremans, A. (1997). Sequence learning in a dual-stimulus setting. *Psychological Research, 60,* 72-86.

Cleeremans, A., & Jiménez, L. (1998). Implicit sequence learning: The truth is in the details. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 323-364). Thousand Oaks, CA: Sage.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120,* 235-253.

Cock, J. J., Berry, D. C., & Buchner, A. (2002). Negative priming and sequence learning. *European Journal of Cognitive Psychology, 14,* 27-48.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cohen, J. D., Dunbar, K., & McClelland, J.L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332-361.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17-30.

Cowan, N. (1995). *Attention and memory: An integrated framework.* Oxford, England: Oxford University Press.

Crabb, B. T., & Dark, V. J. (1999) Perceptual implicit memory requires attentional encoding. *Memory & Cognition, 27,* 267-275.

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 189-202.

Davidson, R. J., Jackson, D. C. & Kalin, N. H. (2000). Emotion, plasticity, context and regulation: Perspectives from affective neuroscience. *Psychological Bulletin, 126,* 890-906.

Dawson, M. E., Schell, A. M., Beers, J. R., & Kelly, A. (1982). Allocation of cognitive resource processing capacity during human autonomic classical conditioning. *Journal of Experimental Psychology: General, 11,* 273-295.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science, 291,* 1803-1806.

Desimone, R. & Duncan, J. (1996). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193-222.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8,* 343-350.

Deutch, J. A., & Deutch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80-90.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92,* 53-78.

Duncan, J. (1999). Converging levels of analysis in the cognitive neuroscience of visual attention. In G. W. Humphreys, J. Duncan, and A. Treisman (Eds.), *Attention, Space, and Action: Studies in Cognitive Neuroscience* (pp. 112-129). New York: Oxford.

Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature, 387*, 808 – 810.

Durlach, P. J., & Rescorla, R. A. (1980). Potentiation rather than overshadowing in flavor-aversion learning: An analysis in terms of within-compound associations. *Journal of Experimental Psychology: Animal Behavior Processes, 6*, 175-187.

Engle, R. W. (2002). Working memory as executive attention. *Current Directions in Psychological Science, 11,* 19-23.

Field, A. P., & Moore, A. C. (2005). Dissociating the effects of Attention and contingency awareness on evaluative conditioning effects in the visual paradigm. *Cognition & Emotion, 19*, 217-243.

Fodor, J. (1983). *The modularity of mind.* Cambridge, MA: MIT/Bradford.

Frensch, P. A. (1998). One concept, multiple meanings: On how to define the concept of implicit learning. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 47-104). Thousand Oaks, CA: Sage.

Frensch, P. A., Buchner, A., & Lin, J. (1994). Implicit learning of unique and ambiguous serial transitions in the presence and absence of a distractor task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 567-584.

Frensch, P. A., Haider, H., Rünger, D., Neugebauer, U., Voigt, S., & Werg, J. (2003). Verbal report of incidentally experienced environmental regularity: The route from implicit learning to verbal expression of what has been learned. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 43-68). Amsterdam & Philadelphia: John Benjamins.

Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research, 61*, 83-98.

Frensch, P. A., & Miner, C. S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition, 22*, 95-110.

Frensch, P. A. & Rünger, D. (2003). Implicit learning. *Current Directions in Psychological Science, 12:1,* 13.

Frensch, P. A., Wenke, D., & Rünger, D. (1999). A secondary tone-counting task suppresses expression of knowledge in the serial reaction task. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25,* 260-274.

Glucksberg, S., & Cowen, G. N. (1970). Memory for nonattended auditory material. *Cognitive Psychology, 1,* 149-156.

Grafton, S. T., Hazeltine, E., & Ivry, R. B. (1998). Abstract and effector-specific representations of motor sequences identified with PET. *Journal of Neurophysiology, 18,* 9420-9428.

Grafton, S. T., Hazeltine, E., & Ivry, R. B. (2002). Motor sequence learning with the nondominant left hand: A PET functional imaging study. *Experimental Brain Research, 146:3,* 369 – 378.

Hayes, N. A., & Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition, 28,* 249-276.

Hazeltine, E., & Ivry, R. B. (2003). Neural structures that support implicit sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 43-68). Amsterdam & Philadelphia: John Benjamins.

Heuer, H., & Schmidtke, V. (1996). Secondary-task effects on sequence learning. *Psychological Research, 59,* 119-133.

Houghton, G., & Tipper, S. P. (1994). A model of inhibitory mechanisms in selective attention. In D. Dagenbach & T. Carr (Eds.), *Inhibitory mechanisms in attention memory and language* (pp. 53-112). San Diego, CA: Academic Press.

Howard, D. V., & Howard, J. H. (2001). When it *does* hurt to try: Adult age differences in the effects of instructions on implicit pattern learning. *Psychonomic Bulletin & Review, 8,* 798-805.

Howard, J. H., Mutter, S. A., & Howard, D. V. (1992). Serial pattern learning by event observation. *Journal of Experimental Psychology: Learning, Memory & Cognition, 18(5),* 1029-39.

Hsiao, A. T., & Reber, A. S. (1998). The role of attention in implicit sequence learning: Exploring the limits of the cognitive unconscious. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 471-494). Thousand Oaks, CA: Sage.

Hsiao, A. T., & Reber, A. S. (2001). The dual-task SRT procedure: Fine-tuning the timing. *Psychonomic Bulletin & Review, 8*, 336-342.

Humphreys, G. W., & Samson, D. (2004). Attention and the frontal lobes. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences,* 3rd Edition (pp. 1091-1104), MIT Press: Cambridge, MA.

Isingrini, M., Vazou, F., & Leroy, P. (1995). Dissociation between implicit and explicit memory tests: Effects of age and divided attention on category exemplar generation and cued recall. *Memory and Cognition, 23*, 462-467.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General, 122*, 139-154.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *Quarterly Journal of Experimental Psychology, 54A*, 1105-1124.

Jiang, Y., & Chun, M. M. (2003). Contextual cueing: Reciprocal influences between attention and implicit learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 277-296). Amsterdam & Philadelphia: John Benjamins.

Jiang, Y., & Leung, A. W. (2005). Implicit learning of ignored visual context. *Psychonomic Bulletin & Review, 12,* 100-106.

Jiménez, L. (Ed.). (2003). *Attention and implicit learning.* Amsterdam & Philadelphia: John Benjamins.

Jiménez, L. (2003). Intention, attention, and consciousness in probabilistic sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 43-68). Amsterdam & Philadelphia: John Benjamins.

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 236-259.

Jiménez, L., & Méndez, C. (2001). Implicit sequence learning with competing explicit cues. *Quarterly Journal of Experimental Psychology, 54A*, 345-369.

Jiménez, L., Méndez, C., & Cleeremans, A. (1996). Direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 948-969.

Jiménez, L., Méndez, C., & Lorda, M. J. (1993). Aprendizaje implícito, atencíon y conciencia en una tarea de tiempo de reaccíon serial [Implicit learning, attention and awareness in a serial, reaction-time task]. *Revista de Psicología General y Aplicada, 46*, 245-255.

Jiménez, L., & Vázquez, G. A. (2005). Sequence learning under dual-task conditions: Alternatives to a resource-based account. *Psychological Research, 69,* 352-368.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Neuroscience, 37*, 43-75.

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science, 6*, 365-369.

Kahneman, D. (1973). *Attention and effort.* Engelwood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 497-509.

Kahneman, D., Treisman, A., & Burkell, J. (1983). The cost of visual filtering. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 510-522.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience, 23*, 315-341.

Keele, S. W., Jennings, P., Jones, S., Caulton, D., & Cohen, A. (1995). On the modularity of sequence representation. *Journal of Motor Behaviour, 27*, 17-30.

Kihlstrom, J. F. (1987). The cognitive unconscious. *Science, 237,* 1445-1452.

Kinder, A., & Shanks, D. R. (2001). Amnesia and the declarative/nondeclarative distinction: A recurrent network model of classification, recognition, and repetition priming. *Journal of Cognitive Neuroscience, 13*, 648-669.

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review, 110*, 728–744.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review, 111*, 880-913.

Lambert, T. (2003). Visual orienting, learning and conscious awareness. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 253-275). Amsterdam & Philadelphia: John Benjamins.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception & Performance, 21,* 451-468.

Lavie, N. (2000). Selective attention and cognitive control: Dissociating attentional functions through different types of load. In: S. Monsell & J. Driver (Eds.). *Attention and Performance XVIII,* 175-194.

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in Cognitive Sciences, 9*, 75-82.

Lavie, N., & de Fockert, J. W. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Perception & Psychophysics, 65,* 202-212.

Lavie, N., & de Fockert, J. W. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review, 12,* 669-74.

Lavie, N., & Fox, E. (2000). The role of perceptual load in negative priming. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1038-1052.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*, 339-354.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics, 56,* 183-197.

Logan, G.D. (1992). Attention and preattention in theories of automaticity. *Journal of Experimental Psychology: General, 105,* 317-39.

Logan, G. D., Taylor, S. E., & Etherton, J. L. (1999). Attention and automaticity: Toward a theoretical integration. *Psychological Research, 62,* 165 – 181.

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behaviour Processes, 28,* 3-26.

Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276-298.

Maruff, P., Danckert, J., Camplin, G., & Currie, J. (1999). Behavioral goals constrain the selection of visual information. *Psychological Science, 10,* 522-525.

Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 350-364.

McCann, R. S., Remington, R. W., & Van Selst, M. (2000). A dual-task investigation of automaticity in visual word processing. *Journal of Experimental Psychology, Human Perception and Performance, 26,* 1352-1370.

McDowall, J., Lustig, A., & Parkin, G. (1995). Indirect learning of event sequences: the effects of divided attention and stimulus continuity. *Canadian Journal of Experimental Psychology, 49,* 415-36.

Merikle, P. (1994). On the futility of attempting to demonstrate null awareness. Commentary on: Shanks, D.R., & St John, M.F. Characteristics of dissociable human learning systems. *Behavioural and Brain Sciences, 17,* 367-447.

Milliken, B., Joordens, S., Merikle, P. M., & Seiffert, A. E. (1998). Selective Attention: A Reevaluation of the Implications of Negative Priming. *Psychological Review, 105,* 203-299.

Mulligan, N. W., & Brown, A. S. (2003). Attention and implicit memory. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 297-334). Amsterdam & Philadelphia: John Benjamins.

Mulligan, N. W., & Hornstein, S. L. (2000). Attention and perceptual priming in the perceptual identification task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 626-637.

Navon, D. (1984). Resources: A theoretical soup stone? *Psychological Review, 91,* 216-234.

Navon, D., & Gopher, D. (1979). On the Economy of the Human-Processing System. *Psychological Review, 86,* 214-255.

Neisser, U. (1976). *Cognition and reality.* San Francisco: Freeman.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence form performance measures. *Cognitive Psychology, 19,* 1-32.

Norman, D. A. (1969). Memory while shadowing. *Quarterly Journal of Experimental Psychology, 21,* 85-93.

Norman, D. A., & Bobrow, D. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44-64.

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science, 9,* 247-255.

Öhman, A. (2002). Automaticity and the amygdala: Nonconscious responses to emotional faces. *Current Directions in Psychological Science, 11,* 62–66.

Parkin, A. J., & Russo, R. (1990). Implicit and explicit and the automatic/effortful distinction. *European Journal of Cognitive Psychology, 2,* 71-80.

Pashler, H. E. (1998). *The psychology of attention.* Cambridge, MA: MIT Press.

Pashler, H. E. (2000). Task switching and multitask performance. In Monsell, S., and Driver, J. (editors). *Attention and Performance* XVIII: Control of mental processes. Cambridge, MA: MIT Press.

Pashler, H. E., Johnston J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology, 52,* 629-651.

Perruchet, P., & Amorim, M. A. (1992). Conscious knowledge and changes in performance in sequence learning: Evidence against dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 785-800.

Perruchet, P., & Vinter, A. (2002). The self-organising consciousness: A framework for implicit learning. In R. M. French and A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, computational and philosophical consensus in the making?* (pp. 41-67). Hove, UK: Psychology Press.

Pessoa, L. (2005). To what extent are emotional visual stimuli processed without attention and awareness? *Current Opinion in Neurobiology, 15,* 188–196.

Pessoa, L., & Ungerleider, L. G. (2004). Top-down mechanisms for working memory and attentional processes. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences,* 3rd Edition (pp. 1091-1104), MIT Press: Cambridge, MA.

Posner, M. I. (1992). Attention as a cognitive and neural system. *Psychological Science, 1,* 11-14.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review, 78*, 391-408.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42.

Rah, S. K., Reber, A. S., & Hsiao, A. T. (2000). Another wrinkle on the dual-task SRT experiment: It's probably not dual task. *Psychonomic Bulletin & Review, 7*, 309-313.

Ratcliff, R., McKoon, G., & Van Zandt, T. (1995). Process dissociation, single-process theories, and recognition memory. *Journal of Experimental Psychology: General, 124*, 352-374.

Reber, A. S. (1967) Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behaviour, 6,* 855-863.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Reber, P. J., & Squire, L. R. (1998). Encapsulation of implicit and explicit memory in sequence learning. *Journal of Cognitive Neuroscience, 10*, 248-263.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20,* 585-594.

Rees, G., Frith, C., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*, *28*, 1616-1619. Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 581-597.

Rees, G., Russell, C., Frith, C. D., & Driver, J. (1999). Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science* 24, 286, 2504 – 2507.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics, 44,* 563-575.

Remillard, G. (2003). Pure perceptual-based sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 581-597.

Remington, R.W., & Folk, C.L. (2001). A dissociation between attention and selection. *Psychological Science, 12*, 511-515.

Rolls, E. T. (2000). Memory systems in the brain. *Annual Review of Psychology, 51,* 599-630.

Rowland, L. A., & Shanks, D. R. (2006a). Sequence learning and selection difficulty. *Journal of Experimental Psychology: Human Perception & Performance, 32,* 287-299.

Rowland, L. A., & Shanks, D. R. (2006b). Attention modulates the learning of multiple contingencies. *Psychonomic Bulletin & Review, 13*, xx-xx.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8,* 101-105.

Schmidtke, V., & Heuer, H. (1997). Task integration as a factor in secondary-task effects on sequence learning. *Psychological Research, 60,* 53-71.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84,* 1-66.

Schvaneveldt, R. W., & Gomez, R. L. (1998). Attention and probabilistic sequence learning. *Psychological Research, 61,* 175-190.

Searle, J. (1992). *The Rediscovery of the Mind.* Cambridge, MA: MIT Press.

Seger, C. A. (1997a). Multiple forms of learning in implicit learning. In M. Stadler & P. Frensch (Eds.) *Handbook of Implicit Learning.* New York: Sage.

Seger, C. A. (1997b). Two forms of sequential implicit learning. *Consciousness and Cognition,* 6, 108-131.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge: Cambridge University Press.

Shanks, D. R. (2003). Attention and awareness in "implicit" sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 11-42). Amsterdam & Philadelphia: John Benjamins.

Shanks, D. R. (2005). Implicit learning. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of cognition* (pp. 202-220). London: Sage.

Shanks, D. R., & Channon, S. (2002). Effects of a secondary task on "implicit" sequence learning: Learning or performance? *Psychological Research, 66,* 99-109.

Shanks, D. R., & Johnstone, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25,* 1435-1451.

Shanks, D. R., Johnstone, T., & Kinder, A. (2002). Modularity and artificial grammar learning. In R. M. French and A. Cleeremans (Eds.), *Implicit learning and*

*consciousness: An empirical, computational and philosophical consensus in the making?* (pp. 93-120). Hove, UK: Psychology Press.

Shanks, D. R., & Perruchet, P. (2002). Dissociation between priming and recognition in the expression of sequential knowledge. *Psychonomic Bulletin & Review, 9*, 362-367.

Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research, 69*, 369-382.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*, 367-447.

Shanks, D. R., Wilkinson, L., & Channon, S. (2003). Relationship between priming and recognition in deterministic and probabilistic sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 248-261.

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. *Psychological Review, 94*, 439-454.

St. John, M. F., & Shanks, D. R. (1997). Implicit learning from an information processing standpoint. In D. C. Berry (Ed.), *How implicit is implicit learning?* (pp.162-194). London: Oxford University Press.

Stadler, M. A. (1995). Role of attention in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 674-685.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18, I643-662.*

Styles, E. A. (1997). *The psychology of attention.* Hove, England: Psychology Press.

Thorndike, E. L. (1931). *Human learning.* New York: Century.

Tipper, S. P., & Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. *Quarterly Journal of Experimental Psychology, 37A*, 591-611.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97-136.

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics, 33*, 527-532.

Treisman, A. M., Vieira, A., & Hayes, A. (1992). Automatic and preattentive processing. *American Journal of Psychology, 105,* 341-362.

Tulving, E. (1995). Organization of memory: Quo vadis?. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences,* 1st Edition (pp. 839-847), MIT Press: Cambridge, MA.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 19,* 301-6.

Tunney, R. J., & Shanks, D. R. (2003). Does opposition logic provide evidence for conscious and unconscious processes in artificial grammar learning? *Consciousness and Cognition, 12*, 201-218.

Unsworth, N., & Engle, R.W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition, 33,* 213-220.

Wickens, C. D. (1984). *Engineering psychology and human performance*. Columbus, OH: Merrill.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 354-369.

Willingham, D. B., Greenberg, A. R., & Thomas, R. C. (1997). Response-to-stimulus interval does not affect implicit motor sequence learning, but does affect performance. *Memory & Cognition, 25,* 534-542.

Willingham, D. B., Wells, L. A., Farrell, J. M., & Stemwedel, M. E. (2000). Implicit motor sequence learning is represented in response locations. *Memory and Cognition, 28*, 366-375.

Yantis, S. (2005). How visual salience wins the battle for awareness. *Nature Neuroscience, 8*, 975-977.

Zaki, S., Nosofsky, R. M., Jessup, N. M., & Unverzagt, F. W. (2003). Categorization and recognition performance of a memory-impaired group: Evidence for single-system models. *Journal of the International Neuropsychological Society, 9*, 394-406.